

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2004, 13:11:17 ; Search time 66 Seconds  
(without alignments)  
1569.204 Million cell updates/sec

Title: US-09-807-148-4  
Perfect score: 971  
Sequence: 1 BPAVFKQFLDGDGWTSRW.....PDNTVEYKIDNSQVSGSLE 180

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02:.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	971	100.0	417	1 CRTIC HUMAN	P27797 homo sapien
2	971	100.0	417	2 AAP36116	Aap36116 homo sapi
3	971	100.0	417	3 CAG33351	Cag33351 homo sapi
4	940	96.8	417	1 CRTIC CRIGR	Q8K3H7 cricetus
5	933	96.1	416	1 CRTIC RAT	P18418 rattus norv
6	933	96.1	416	2 AAH62395	Aah62395 rattus no
7	931	95.9	418	1 CRTIC RABIT	P15253 oryctolagus
8	930	95.8	416	1 CRTIC MOUSE	P14211 mus muscu
9	930	95.8	416	2 BAC35852	Bac35852 mus muscu
10	930	95.8	417	1 CRT1 BOVIN	P52193 bos taurus
11	874	90.0	404	2 Q6EE32	Q6ee32 gallus gall
12	843	86.8	413	2 Q7ZW08	Q7zw08 xenopus lae
13	843	86.8	418	2 Q7ZXV3	Q7zxv3 xenopus lae
14	835	86.0	411	2 Q91710	Q91710 xenopus lae
15	820	84.4	416	2 Q6NV76	Q6nv76 xenopus tro
16	820	84.4	415	2 AAH67917	Aah67917 xenopus t
17	817	84.1	419	2 Q6PF78	Q6pf78 brachydanio
18	817	84.1	419	2 AAH57469	Aah57469 brachydan
19	804	82.8	405	2 Q26268	Q26268 aplysia cal
20	797	82.1	343	2 Q91711	Q91711 xenopus lae
21	796	82.0	419	2 Q98984	Q98984 rana rugosa
22	785	80.8	417	2 Q9PU01	Q9pu01 brachydanio
23	785	80.8	418	2 Q6NV30	Q6nv30 brachydanio
24	785	80.8	418	2 Q6PE26	Q6pe26 brachydanio
25	785	80.8	418	2 AAH58314	Aah58314 brachydan
26	785	80.8	418	2 AAH68336	Aah68336 brachydan
27	773	79.6	407	2 Q6WSP9	Q6wsp9 dermacentor
28	773	79.6	407	2 Q86G72	Q86g72 dermacentor
29	773	79.6	407	2 AAQ18697	Aaq18697 dermacent
30	769	79.2	423	2 Q7SZM3	Q7szm3 ictalurus p
31	763	78.6	417	2 Q802X3	Q802x3 brachydanio

RESULT 1				
CRTIC_HUMAN				
ID	CRTIC_HUMAN	STANDARD;	PRT;	417 AA.
AC	P27797;			
DT	01-AUG-1992 (Rel. 23, Created)			
DT	01-AUG-1992 (Rel. 23, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	Calreticulin precursor [CRP55] (Calregulin) (HACBP) (ERP60) (grp60).			
GN	Name=CALR; Synonyms=CRIC;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92013129; PubMed=1919005;			
RA	Rokeach L.A., Haselby J.A., Meilof J.F., Smeenk R.J., Unnasch T.R.,			
RA	Greene B.M., Hoch S.O.;			
RT	"Characterization of the autoantigen calreticulin."			
RL	J. Immunol. 147:3031-3039(1991).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=90237213; PubMed=2332496;			
RA	McCaulliffe D.P., Lux F.A., Lieu T.S., Sanz I., Hanke J., Newkirk M.M.,			
RA	Bachinski L.L., Itoh Y., Siciliano M.J., Reichlin M., Sontheimer R.D.,			
RA	Capra J.D.;			
RT	"Molecular cloning, expression, and chromosome 19 localization of a			
RT	human Ro/SS-A autoantigen."			
RL	J. Clin. Invest. 85:1379-1391(1990).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=92129342; PubMed=1733953;			
RA	McCaulliffe D.P., Yang Y.S., Wilson J., Sontheimer R.D., Capra J.D.;			
RT	"The 5'-flanking region of the human calreticulin gene shares homology			
RT	with the human GRP78, GRP94, and protein disulfide isomerase			
RT	promoters."			
RL	J. Biol. Chem. 267:2557-2562(1992).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Liu J., Peng X., Yuan J., Qiang B.;			
RA	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RX	PubMed=15057824; DOI=10.1038/nature02399;			
RA	Grimwood J., Gordon L.A., Olsen A.S., Terry A., Schmutz J., Tran-Gyamfi M.,			
RA	Lamerdin J.E., Hellsten U., Goodstein D., Couronne O., Branscomb E.,			
RA	Aerts A., Altherr M., Ashworth L., Bajorek E., Black S., Branscomb E.,			
RA	Caenepeel S., Carrano A.V., Caille C., Chan Y.M., Christensen M.,			
RA	Cleland C.A., Copeland A., Dalin E., Dehal P., Denys M., Deter J.C.,			
RA	Escobar J., Flowers D., Fotopoulos D., Garcia C., Georgescu A.M.,			
RA	Glavina T., Gomez M., Gonzales E., Groza M., Hammon N., Hawkins T.,			
RA	Raydu L., Ho I., Huang W., Israni S., Jett J., Kacher K., Kimball H.,			
RA	Kobayashi A., Larionov V., Leem S.-H., Lopez F., Lou Y., Lowry S.,			
RA	Malfatti S., Martinez D., McCready P.M., Medina C., Morgan J.,			

Q6dl13 brachydanio  
Q6wsq2 rhinicephal  
Aaq18694 rhiniceph  
Q8mwp3 boophilus m  
Q16893 amblyomma a  
Q6wsq1 haemaphysal  
Aaq18695 haemaphys  
Q6wsq0 ixodes scap  
Q6wsq0 ixodes sc  
Aaq18696 ixodes sc  
Q6ud79 oncorhynchu  
Aa17084 oncorhynch  
Q6u6s0 strongyloce  
Q7pqk3 anopheles g  
Q8wr36 anopheles g

32 763 78.6 417 2 Q6dl13  
33 762 78.5 411 2 Q6wsq2  
34 762 78.5 411 2 Aaq18694  
35 761 78.4 411 2 Q8mwp3  
36 759 78.2 410 2 Q16893  
37 756 77.9 410 2 Q6wsq1  
38 756 77.9 410 2 Aaq18695  
39 755.5 77.8 413 2 Q6wsq0  
40 755.5 77.8 413 2 Aaq18696  
41 746 76.8 419 2 Q6ud79  
42 746 76.8 419 2 Aa17084  
43 731 75.3 421 2 Q9u6s0  
44 701 72.2 406 2 Q7pqk3  
45 701 72.2 406 2 Q8wr36

RA Nelson K., Nolan M., Ovcharenko I., Pitluck S., Pollard M.,  
RA Popkie A.P., Predki P., Quan G., Ramirez L., Rash S., Retterer J.,  
RA Rodriguez A., Rogers S., Salazar A., Salazar A., She X., Smith D.,  
RA Slezak T., Solovyev V., Thayer M., Tice H., Tsai M., Ustaszewska A.,  
RA Vo N., Wagner M., Wheeler J., Wu K., Xie G., Yang J., Dubchak I.,  
RA Furey T.S., DeJong P., Dickson M., Gordon D., Eichler E.E.,  
RA Pennacchio L.A., Richardson P., Stubbs L., Rokhsar D.S., Myers R.M.,  
RA Rubin E.M., Lucas S.M.;  
RL "The DNA sequence and biology of human chromosome 19";  
RL Nature 428:529-535(2004).  
RN [6]  
RN SEQUENCE FROM N.A.  
RC TISSUE=Eye, Pancreas, and Skin;  
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettaman M., Madan A., Rodriguez S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Butterfield Y.S.N., Krzyzinski M.I., Skalska U., Smalilus D.E.,  
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
and mouse cDNA sequences";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [7]  
RN SEQUENCE OF 18-41.  
RX PubMed=3260607;  
RA Lieu T.-S., Newkirk M.M., Capra J.D., Sontheimer R.D.;  
RT "Molecular characterization of human Ro/SS-A antigen. Amino terminal  
sequence of the protein moiety of human Ro/SS-A antigen and  
immunological activity of a corresponding synthetic peptide";  
RL J. Clin. Invest. 82:96-101(1988).  
RN [8]  
RN SEQUENCE OF 18-36.  
RX MEDLINE=92002034; PubMed=19111778;  
RA Rojiani M.V., Finlay B.B., Gray V., Dedhar S.;  
RT "In vitro interaction of a polypeptide homologous to human Ro/SS-A  
antigen (calreticulin) with a highly conserved amino acid sequence in  
the cytoplasmic domain of integrin alpha subunits";  
RL Biochemistry 30:9859-9866(1991).  
RN [9]  
RN SEQUENCE OF 18-32.  
RX MEDLINE=90380058; PubMed=2400400;  
RA Krause K.H., Simmerman H.K.B., Jones L.R., Campbell K.P.;  
RT "Sequence similarity of calreticulin with a Ca2(+)-binding protein  
that co-purifies with an Ins(1,4,5)P3-sensitive Ca2+ store in HL-60  
cells";  
RL Biochem. J. 270:545-548(1990).  
RN [10]  
RN SEQUENCE OF 18-28.  
RC TISSUE=Liver;  
RX MEDLINE=93162045; PubMed=1286669;  
RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,  
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,  
RA Appel R.D., Hughes G.J.;  
RT "Human liver protein map: a reference database established by  
microsequencing and gel comparison";  
RL Electrophoresis 13:992-1001(1992).  
RN [11]  
RN PARTIAL SEQUENCE OF 25-34; 56-62; 208-221 AND 273-278.  
RC TISSUE=Keratinocytes;  
RX MEDLINE=93162043; PubMed=1286667;  
RA Rasmussen H.H., van Damme J., Puype M., Gesser B., Celis J.E.,  
RA Vandekerckhove J.;  
RT "Microsequences of 145 proteins recorded in the two-dimensional gel  
protein database of normal human epidermal keratinocytes";  
RL Electrophoresis 13:960-969(1992).  
RN [12]  
RN SEQUENCE OF 18-27, AND SUBCELLULAR LOCATION.  
RX PubMed=8418194;  
RA Dupuis M., Schaefer E., Krause K.-H., Tschopp J.;  
RT "The calcium-binding protein calreticulin is a major constituent of  
lytic granules in cytolytic T lymphocytes";  
RL J. Exp. Med. 177:1-7(1993).  
RN [13]  
RN SEQUENCE OF 18-26.  
RC TISSUE=Colon carcinoma;  
RX MEDLINE=97295306; PubMed=9150948;  
RA Ji H., Reid G.E., Moritz R.L., Eddes J.S., Burgess A.W., Simpson R.J.;  
RT "A two-dimensional gel database of human colon carcinoma proteins";  
RL Electrophoresis 18:605-613(1997).  
RN [14]  
RN FUNCTION.  
RX MEDLINE=95181473; PubMed=7876246;  
RA Nauseef W.M., McCormick S.J., Clark R.A.;  
RT "Calreticulin functions as a molecular chaperone in the biosynthesis  
of myeloperoxidase";  
RL J. Biol. Chem. 270:4741-4747(1995).  
RN [15]  
RN SUBCELLULAR LOCATION.  
RX MEDLINE=99287886; PubMed=10358038;  
RA Atosa F.A., de Jesus O., Porto G., Carmo A.M., de Sousa M.;  
RT "Calreticulin is expressed on the cell surface of activated human  
peripheral blood T lymphocytes in association with major  
histocompatibility complex class I molecules";  
RL J. Biol. Chem. 274:16917-16922(1999).  
RN [16]  
RN FUNCTION, INTERACTION WITH NR3C1, SUBCELLULAR LOCATION, AND MASS  
SPECTROMETRY.  
RX MEDLINE=21067064; PubMed=11149926;  
RA Holaska J.M., Black B.E., Love D.C., Hanover J.A., Leszyk J.,  
RA Paschal B.M.;  
RT "Calreticulin is a receptor for nuclear export";  
RL J. Cell Biol. 152:127-140(2001).  
RN [17]  
RN PARTIAL SEQUENCE, MASS SPECTROMETRY, AND DISULFIDE BOND.  
RC TISSUE=Placenta;  
RX MEDLINE=21223009; PubMed=11322874;  
RA Hoejrup P., Roepstorff P., Houen G.;  
RT "Human placental calreticulin characterization of domain structure and  
post-translational modifications";  
RL Eur. J. Biochem. 268:2558-2565(2001).  
CC -/- FUNCTION: Molecular calcium binding chaperone promoting folding,  
calreticulin/calnexin cycle. This lectin interacts transiently  
with almost all of the monoglucosylated glycoproteins that are  
synthesized in the ER. Interacts with the DNA-binding domain of  
NR3C1 and mediates its nuclear export.  
CC -/- SUBUNIT: Monomer. Interacts with GRP58 (ERP57) and with NR3C1 (By  
similarity).  
CC -/- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen. Also found in  
cell surface (T cells), cytosol and extracellular matrix.  
CC Associated with the lytic granules in the cytolytic T-lymphocytes.  
CC -/- DOMAIN: Can be divided into a N-terminal globular domain, a  
proline-rich P-domain forming an elongated arm-like structure and  
a C-terminal acidic domain. The P-domain binds one molecule of  
calcium with high affinity, whereas the acidic C-domain binds  
multiple calcium ions with low affinity.  
CC -/- DOMAIN: The interaction with glycans occurs through a binding site  
in the globular lectin domain.  
CC -/- DOMAIN: The zinc binding sites are localized to the N-domain.  
CC -/- DOMAIN: Associates with ERP57 through the tip of the extended arm  
formed by the P-domain.  
CC -/- MASS SPECTROMETRY: MW=46879; METHOD=MALDI; RANGE=18-417;  
CC NOTE=Ref.16.  
CC -/- SIMILARITY: Belongs to the calreticulin family.

CC -!- CAUTION: Was originally (Ref.2) thought to be the 52 kDa Ro  
CC autoantigen.

Query Match 100.0%; Score 971; DB 1; Length 417;  
Best Local Similarity 100.0%; Pred. No. 6.4e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EPAYVFEQFLDGDGWTSRWIESKHSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVFEQFLDGDGWTSRWIESKHSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 77  
QY 61 SASPEFSPNKGQTLVQVFTVKHEQNIIDCGGYVKLFPNSLDQTDHGDSEYNIMFGPDIC 120  
DB 78 SASPEFSPNKGQTLVQVFTVKHEQNIIDCGGYVKLFPNSLDQTDHGDSEYNIMFGPDIC 137  
QY 121 GPGTKKVHVIENYKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 180  
DB 138 GPGTKKVHVIENYKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 197

## RESULT 2

AAP36116 PRELIMINARY; PRT; 417 AA.  
ID AAP36116  
AC AAP36116;  
DT 02-MAR-2004 (TrEMBLrel. 27, Created)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TrEMBLrel. 27, Last annotation update)  
DE Calreticulin.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
[1]  
RN SEQUENCE FROM N.A.  
RP Kalline N., Chen X., Rolfes A., Halleck A., Hines L., Eisenstein S.,  
RA Koundinya M., Raphael J., Moreira D., Kelley T., Labaer J., Lin Y.,  
RA Bhelan M., Farmer A.  
RT "Cloning of human full-length cDNAs in BD Creator(TM) System Donor  
RT vector."  
RL Submitted (MAY-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BT007448; AAP36116.1; -  
SQ SEQUENCE 417 AA; 48141 MW; BC37C3C0F1054FB2 CRC64;

Query Match 100.0%; Score 971; DB 2; Length 417;  
Best Local Similarity 100.0%; Pred. No. 6.4e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EPAYVFEQFLDGDGWTSRWIESKHSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVFEQFLDGDGWTSRWIESKHSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 77  
QY 61 SASPEFSPNKGQTLVQVFTVKHEQNIIDCGGYVKLFPNSLDQTDHGDSEYNIMFGPDIC 120  
DB 78 SASPEFSPNKGQTLVQVFTVKHEQNIIDCGGYVKLFPNSLDQTDHGDSEYNIMFGPDIC 137  
QY 121 GPGTKKVHVIENYKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 180  
DB 138 GPGTKKVHVIENYKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 197

## RESULT 3

CAG33351 PRELIMINARY; PRT; 417 AA.  
ID CAG33351  
AC CAG33351;  
DT 01-JUN-2004 (TrEMBLrel. 27, Created)  
DT 01-JUN-2004 (TrEMBLrel. 27, Last sequence update)  
DT 01-JUN-2004 (TrEMBLrel. 27, Last annotation update)  
DE CALR protein.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;

RN SEQUENCE FROM N.A.  
RP Ebert L., Schick M., Neubert P., Schatten R., Henze S., Korn B.;  
RA "Cloning of human full open reading frames in Gateway(TM) system entry  
RT vector (pDONR201).";  
RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.  
DR EMBL; CR457070; CAG33351.1; -  
SQ SEQUENCE 417 AA; 48141 MW; BC37C3C0F1054FB2 CRC64;  
Query Match 100.0%; Score 971; DB 2; Length 417;  
Best Local Similarity 100.0%; Pred. No. 6.4e-79;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EPAYVFEQFLDGDGWTSRWIESKHSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVFEQFLDGDGWTSRWIESKHSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 77  
QY 61 SASPEFSPNKGQTLVQVFTVKHEQNIIDCGGYVKLFPNSLDQTDHGDSEYNIMFGPDIC 120  
DB 78 SASPEFSPNKGQTLVQVFTVKHEQNIIDCGGYVKLFPNSLDQTDHGDSEYNIMFGPDIC 137  
QY 121 GPGTKKVHVIENYKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 180  
DB 138 GPGTKKVHVIENYKGNVINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 197

## RESULT 4

CRIC\_CRICR STANDARD; PRT; 417 AA.  
ID CRIC\_CRICR  
AC Q8K3H7;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERP60).  
GN Name=CALR;  
OS Cricetulus griseus (Chinese hamster).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
OC Cricetulus.  
OX NCBI\_TaxID=10029;  
[1]  
RN SEQUENCE FROM N.A.  
RP Chung J.Y., Lee G.M.;  
RA Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.  
RL FUNCTION: Molecular calcium binding chaperone promoting folding,  
CC oligomeric assembly and quality control in the ER via the  
CC calreticulin/calnexin cycle. This lectin interacts transiently  
CC with almost all of the monoglucosylated glycoproteins that are  
CC synthesized in the ER. Interacts with the DNA-binding domain of  
CC NR3C1 and mediates its nuclear export (By similarity).  
CC -!- SUBUNIT: Monomer. Interacts with GRP58 (ERP57) and with NR3C1 (By  
CC similarity).  
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.  
CC -!- DOMAIN: Can be divided into a N-terminal globular domain, a  
CC proline-rich P-domain forming an elongated arm-like structure and  
CC a C-terminal acidic domain. The P-domain binds one molecule of  
CC calcium with high affinity, whereas the acidic C-domain binds  
CC multiple calcium ions with low affinity (By similarity).  
CC -!- DOMAIN: The interaction with glycans occurs through a binding site  
CC in the globular lectin domain (By similarity).  
CC -!- DOMAIN: The zinc binding sites are localized to the N-domain (By  
CC similarity).  
CC -!- DOMAIN: Associates with ERP57 through the tip of the extended arm  
CC formed by the P-domain (By similarity).  
CC -!- SIMILARITY: Belongs to the calreticulin family.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).



CC -I- SUBUNIT: Monomer. Interacts with GRP58 (ERP57) and with NR3C1 (By  
CC similarity).

CC -I- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.

CC -I- TISSUE SPECIFICITY: Predentin and odontoblast.

CC -I- DOMAIN: Can be divided into a N-terminal globular domain, a  
CC proline-rich p-domain forming an elongated arm-like structure and  
CC a C-terminal acidic domain. The p-domain binds one molecule of  
CC calcium with high affinity, whereas the acidic C-domain binds  
CC multiple calcium ions with low affinity (By similarity).

CC -I- DOMAIN: The interaction with glycans occurs through a binding site  
CC in the globular lectin domain (By similarity).

CC -I- DOMAIN: The zinc binding sites are localized to the N-domain (By  
CC similarity).

CC -I- DOMAIN: Associates with ERP57 through the tip of the extended arm  
CC formed by the p-domain (By similarity).

CC -I- SIMILARITY: Belongs to the calreticulin family.

CC -I- CAUTION: Was originally (Ref.2) thought to be D-beta-  
CC hydroxybutyrate dehydrogenase.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL; D78308; BAA11345.1; --  
DR EMBL; X53363; CAA37446.1; --  
DR EMBL; X13702; CAA31987.1; ALT\_SEQ.  
DR EMBL; X79327; CAA55890.1; --  
DR PIR; JH0819; JH0819.  
DR PDB; 1HHN; NMR; A=205-305.  
DR PDB; 1K91; NMR; A=237-273.  
DR PDB; 1K9C; NMR; A=205-278.  
DR RGD; 620288; Calr.  
DR InterPro; IPR001580; Calret/calnex.  
DR InterPro; IPR009169; Calreticulin.  
DR InterPro; IPR009033; Calret/calnex\_P.  
DR InterPro; IPR008985; ConA like lec\_gl.  
DR InterPro; IPR000886; ER target\_S.  
DR Pfam; PF00262; Calreticulin; 1.  
DR PIRSF; PIRSF002356; Calreticulin; 1.  
DR PRINTS; PR00626; CALRETTICULIN.  
DR ProDom; PD001866; Calreticulin; 1.  
DR PROSITE; PS00803; CALRETTICULIN\_1; 1.  
DR PROSITE; PS00804; CALRETTICULIN\_2; 1.  
DR PROSITE; PS00805; CALRETTICULIN\_REPEAT; 3.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
DR 3D-structure; Calcium-binding; Chaperone; Direct protein sequencing;  
KW Endoplasmic reticulum; Lectin; Metal-binding; Repeat; Signal; Zinc.  
FT SIGNAL 1 17  
FT CHAIN 18 416 Calreticulin.  
FT DOMAIN 18 197 N-domain.  
FT DOMAIN 198 308 P-domain.  
FT DOMAIN 309 416 C-domain.  
FT DOMAIN 191 255 4 X approximate repeats.  
FT REPEAT 191 202 1-1.  
FT REPEAT 210 221 1-2.  
FT REPEAT 227 238 1-3.  
FT REPEAT 244 255 1-4.  
FT DOMAIN 259 297 3 X approximate repeats.  
FT REPEAT 259 269 2-1.  
FT REPEAT 273 283 2-2.  
FT REPEAT 287 297 2-3.  
FT DOMAIN 351 407 Asp/Glu/Lys-rich.  
FT DISULFID 105 137 By similarity.  
FT SITE 413 416 Prevent secretion from ER.  
FT STRAND 208 208  
FT TURN 212 213  
FT TURN 221 222  
FT TURN 224 225  
FT TURN 235 236

FT STRAND 241 243  
FT TURN 245 246  
FT TURN 251 252  
FT HELIX 255 258  
FT STRAND 265 267  
FT TURN 269 270  
FT STRAND 280 281  
FT STRAND 294 294  
FT TURN 297 298  
SQ SEQUENCE 416 AA; 47995 MW; 2B6713CED31A2970 CRC64;

Query Match 96.1%; Score 933; DB 1; Length 416;  
Best Local Similarity 94.4%; Pred. No. 1.7e-75;  
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

OY 1 EPAYVFEQFDGQWTSRWIESKHSDFGKFLVSSGKFGDEBKDKGLQTSQDARFYAL 60  
DB 18 DPAYVFEQFDGDAWNRWVESKHSDFGKFLVSSGKFGDEBKDKGLQTSQDARFYAL 77  
OY 61 SASPEPSNKGQTLVVQFTVKHEQNIDCGGKGVKLPNSLDQDTHGDSSEYNINFGPDIC 120  
DB 78 SARPEPSNKGQTLVVQFTVKHEQNIDCGGKGVKLPFGGLDQKMHGDSSEYNINFGPDIC 137  
OY 121 GPGTKKHVHIFPNYKGNVLINKDIRCKDDSFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 180  
DB 138 GPGTKKHVHIFPNYKGNVLINKDIRCKDDSFTHLYTLIVRPDNTYEVKIDNSQVESGSLE 197

RESULT 6  
AAH62395 PRELIMINARY; PRT; 416 AA.  
ID AAH62395  
AC AAH62395; (Created)  
DT 02-MAR-2004 (TRENBLrel. 27, Last sequence update)  
DT 02-MAR-2004 (TRENBLrel. 27, Last annotation update)  
DE Calreticulin.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RX MEDLINE=22388257; PubMed=12477932;  
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
RA Krzywinski M.I., Skalska J., Smallos D.E., Schnerch A., Schein J.E.,  
RA Jones S.J., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Prostate;  
RA Strausberg R.;  
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; BC062395; AAH62395.1; --  
SQ SEQUENCE 416 AA; 47995 MW; 2E6713CED31A2970 CRC64;

Query Match 96.1%; Score 933; DB 2; Length 416;

Best Local Similarity 94.4%; Pred. No. 1.7e-75;  
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPVAYKEQFLDGDGHTSWISKHSDFGKFLVSGKTYGDBEYKDLQISQDARFYAL 60  
Db DPAYKEQFLDGDGHTSWISKHSDFGKFLVSGKTYGDBEYKDLQISQDARFYAL 77

QY 61 SASFEFNSKQTLVQVFTVKEHQNDICGGYVVKLPNPSLDQTDVHGDSYNNMFGPDIC 120  
Db SARFEFNSKQTLVQVFTVKEHQNDICGGYVVKLPNPSLDQTDVHGDSYNNMFGPDIC 137

QY 121 GPOTKXVHVFYNGKQVNLINKDIRCKDDEFFHLVTLVIRPNTYEVKIDNSQVBSGSLE 180  
Db GPOTKXVHVFYNGKQVNLINKDIRCKDDEFFHLVTLVIRPNTYEVKIDNSQVBSGSLE 197

RESULT 7  
CRTC-RABIT STANDARD; PRT; 418 AA.

AC P15253;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 05-JUL-2004 (Rel. 44, Last annotation update)  
DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (BRp60).  
GN Name=CAUR;  
OS Oryctolagus cuniculus (Rabbit).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
OX NCBI\_TaxID=9986;  
RN [1] \_SEQUENCE FROM N.A.  
RP TISSUE=slow-twitch skeletal muscle;  
RC MEDLINE=90094320; PubMed=26000080;  
RA Fliegel L., Burns K., MacLennan D.H., Reithmeier R.A.F., Michalak M.;  
RT "Molecular cloning of the high affinity calcium-binding protein  
RT (calreticulin) of skeletal muscle sarcoplasmic reticulum";  
RL J. Biol. Chem. 264:21522-21528(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=fast-twitch skeletal muscle;  
RX MEDLINE=91282795; PubMed=2059224;  
RA Fliegel L., Michalak M.;  
RT "Fast-twitch and slow-twitch skeletal muscles express the same isoform  
RT of calreticulin";  
RL Biochem. Biophys. Res. Commun. 177:979-984(1991).  
RN [3]  
RP SEQUENCE OF 18-36.  
RX MEDLINE=91054414; PubMed=2241926;  
RA Treves S., de Mattei M., Lanfredi M., Villa A., Green N.M.,  
RA MacLennan D.H., Meldolesi J., Pozzan T.;  
RT "Calreticulin is a candidate for a calsequestrin-like function in  
RT Ca2(+)-storage compartments (calciosomes) of liver and brain";  
RL Biochem. J. 271:473-480(1990).  
RN [4]  
RP SEQUENCE OF 18-46.  
RX MEDLINE=91201375; PubMed=2016321;  
RA Milner R.E., Baksh S., Shenanko C., Carpenter M.R., Smillie L.,  
RA Vance J.E., Opas M., Michalak M.;  
RT "Calreticulin, and not calsequestrin, is the major calcium binding  
RT protein of smooth muscle sarcoplasmic reticulum and liver endoplasmic  
RT reticulum";  
RL J. Biol. Chem. 266:7155-7165(1991).  
RN [5]  
RP PARTIAL SEQUENCE.  
RC TISSUE=Lung;  
RX MEDLINE=92002038; PubMed=1911780;  
RA Guan S., Falick A.M., Williams D.E., Cashman J.R.;  
RT "Evidence for complex formation between rabbit lung flavin-containing  
RT monooxygenase and calreticulin";  
RL Biochemistry 30:9892-9900(1991).  
RN [6]  
RP FUNCTION.  
RX MEDLINE=20050566; PubMed=10581245;

RA Saito Y., Ihara Y., Leach M.R., Cohen-Doyle M.F., Williams D.B.;  
RT "Calreticulin functions in vitro as a molecular chaperone for both  
RT glycosylated and non-glycosylated proteins.";  
RL EMBL J. 18:6718-6729(1999).  
RN [7]  
RP ZINC-BINDING DOMAIN.  
RX PubMed=8521965; DOI=10.1016/0014-5793(95)01246-4;  
RA Baksh S., Spamer C., Heilmann C., Michalak M.;  
RT "Identification of the Zn2+ binding region in calreticulin.";  
RL FEBS Lett. 376:53-57(1995).  
RN [8]  
RP MUTAGENESIS OF HIS-170.  
RX PubMed=14522955; DOI=10.1074/jbc.M309497200;  
RA Guo L., Groenendyk J., Papp S., Dabrowska M., Knoblauch B., Kay C.,  
RA Parker J.M., Opas M., Michalak M.;  
RT "Identification of an N-domain histidine essential for chaperone  
RT function in calreticulin";  
RL J. Biol. Chem. 278:50645-50653(2003).  
CC -I- FUNCTION: Molecular calcium binding chaperone promoting folding,  
CC oligomeric assembly and quality control in the ER via the  
CC calreticulin/calnexin cycle. This lectin interacts transiently  
CC with almost all of the monoglucosylated glycoproteins that are  
CC synthesized in the ER. Interacts with the DNA-binding domain of  
CC NR3C1 and mediates its nuclear export (By similarity).  
CC -I- SUBUNIT: Monomer. Interacts with GRP58 (ERP57) and with NR3C1 (By  
CC similarity).  
CC -I- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.  
CC -I- DOMAIN: Can be divided into a N-terminal globular domain, a  
CC proline-rich P-domain forming an elongated arm-like structure and  
CC a C-terminal acidic domain. The P-domain binds one molecule of  
CC calcium with high affinity, whereas the acidic C-domain binds  
CC multiple calcium ions with low affinity (By similarity).  
CC -I- DOMAIN: The interaction with glycans occurs through a binding site  
CC in the globular lectin domain (By similarity).  
CC -I- DOMAIN: The zinc binding sites are localized to the N-domain.  
CC -I- DOMAIN: Associates with ERP57 through the tip of the extended arm  
CC formed by the P-domain (By similarity).  
CC -I- SIMILARITY: Belongs to the calreticulin family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; J051138; AAA31188.1; -.  
DR PIR; A34154; A34154.  
DR PIR; C33208; C33208.  
DR PIR; S13046; S13046.  
DR HSP; P18418; I1891.  
DR InterPro; IPR001580; Calret/calnex.  
DR InterPro; IPR009169; Calreticulin.  
DR InterPro; IPR009033; Calret calnex P.  
DR InterPro; IPR008985; ConA like lec\_g1.  
DR Pfam; PF00362; Calreticulin\_1.  
DR PRINTS; PIRSF002356; Calreticulin; 1.  
DR PRODOM; PD001866; Calreticulin; 1.  
DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
DR PROSITE; PS00014; ER\_TARGET; 1.  
KW Calcium-binding; Chaperone; Direct protein sequencing;  
KW Endoplasmic reticulum; Lectin; Metal-binding; Repeat; Signal; Zinc.  
FT SIGNAL 1 17  
FT CHAIN 18 418 Calreticulin.  
FT DOMAIN 18 197 N-domain.  
FT DOMAIN 198 308 P-domain.  
FT DOMAIN 309 418 C-domain.  
FT DOMAIN 191 255 4 X approximate repeats.

FT REPEAT 191 202 1-1.  
 FT REPEAT 210 221 1-2.  
 FT REPEAT 227 238 1-3.  
 FT REPEAT 244 255 1-4.  
 FT DOMAIN 259 297 3 X approximate repeats.  
 FT REPEAT 259 269 2-1.  
 FT REPEAT 273 283 2-2.  
 FT REPEAT 287 297 2-3.  
 FT DOMAIN 351 408 Asp/Glu/Lys-rich.  
 FT DISULFID 105 137 By similarity.  
 FT SITE 415 418 Prevent secretion from ER.  
 FT VARIANT 35 35 E -> D.  
 FT MUTAGEN 170 170 H->A: Loss of activity.  
 FT CONFLICT 90 90 P -> T (in Ref. 5).  
 SQ SEQUENCE 418 AA; 48275 MW; B6082B689DCV763A6 CRC64;

Query Match 95.9%; Score 931; DB 1; Length 418;  
 Best Local Similarity 95.6%; Pred. No. 2.5e-75;  
 Matches 172; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 EPVYFKQQLDGDGWTSRVIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSODARFYAL 60  
 DB 18 EPVYFKQQLDGDGWTSRVIESKHKSDFGKFLVSSGKFGYDEKDKGLQTSODARFYAL 77

QY 61 SASFEPSPKNGQTLVQVFTVKHEQNTDCGGYVKLPNSLDQDMHGDSEYNIMFGPDIC 120  
 DB 78 SASFEPSPKNGQPLVQVFTVKHEQNTDCGGYVKLPAGLDQDMHGDSEYNIMFGPDIC 137

QY 121 GPGTKKHVIFNYKGNVLINKDIRKDBFTHLYLIVRPDNTYEVKIDNSQVSGSLE 180  
 DB 138 GPGTKKHVIFNYKGNVLINKDIRKDBFTHLYLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 8

CRTC\_MOUSE STANDARD; PRT; 416 AA.

ID CRTC\_MOUSE  
 AC P14211;  
 DT 01-JAN-1990 (Rel. 13, Created)  
 DT 01-JAN-1990 (Rel. 13, Last sequence update)  
 DT 01-OCT-2004 (Rel. 45, Last annotation update)  
 DE Calreticulin precursor (CRP55) (Calregulin) (HACBP) (ERP60).  
 GN Name=Calr;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 18-48 AND 129-161.  
 RC STRAIN=BALB/c; TISSUE=Liver;  
 RX MEDLINE=90059955; PubMed=2583110;  
 RA Smith M.J., Koch G.L.E.;  
 RT "Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a major calcium binding ER/SR protein.";  
 RL EMBO J. 8:3581-3586(1989).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93013037; PubMed=1398135;  
 RA Mazzarella R.A., Gold P., Cunningham M., Green M.;  
 RT "Determination of the sequence of an expressible cDNA clone encoding ERp60/calregulin by the use of a novel nested set method.";  
 RL Gene 120:217-225(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN=C57BL/6J; TISSUE=Brain;  
 RC MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi J., Bono H., Kondo S.,  
 RA Nakaide I., Osato N., Saito R., Suzuki H., Yananaka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gojobori T.,  
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schrim L.M., Karapin A., Matsuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Bradt D., Brusic V., Chothia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Gustincich S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagashima T., Numata K., Okido T., Pavan W.J., Perteau G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,  
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,  
 RA Sandelin A., Schneider C., Sempile C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wainstedt C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyshaw-Boris A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=FVB/N-3; TISSUE=Mammary gland;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Klausner R.D., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Strausner R.D., Collins F.S., Wagner L., Shennen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Sanchez A.,  
 RA Fahey J., Helton E., Kettman M., Madan A.C., Rodrigues S., Sanchez A.,  
 RA Whiting M., Madan A., Young J.W., Green E.D., Dickson M.C.,  
 RA Blakesley R.W., Touchman J.W., Shevchenko Y., Bouffard G.G.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallos D.E.,  
 RA Schnerth A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RP SEQUENCE OF 18-38.  
 RC TISSUE=Fibroblast;  
 RX MEDLINE=95009907; PubMed=7523108;  
 RA Merrick B.A., Patterson R.M., Wichter L.L., He C., Selkirk J.K.;  
 RT "Separation and sequencing of familial and novel murine proteins using preparative two-dimensional gel electrophoresis.";  
 RL Electrophoresis 15:735-745(1994).  
 RN [6]  
 RP SUBCELLULAR LOCATION.  
 RX PubMed=8418194;  
 RA Dupuis M., Schaefer E., Krause K.-H., Tschopp J.;  
 RT "The calcium-binding protein calreticulin is a major constituent of lytic granules in cytolytic T lymphocytes.";  
 RL J. Exp. Med. 177:1-7(1993).  
 CC -!- FUNCTION: Molecular calcium binding chaperone promoting folding, oligomeric assembly and quality control in the ER via the calreticulin/calnexin cycle. This lectin interacts transiently with almost all of the monoglucosylated glycoproteins that are synthesized in the ER. Interacts with the DNA-binding domain of NR3C1 and mediates its nuclear export (by similarity).  
 CC -!- SUBUNIT: Monomer. Interacts with GRP58 (ERP57) and with NR3C1 (by similarity).  
 CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen. Associated with the lytic granules in the cytolytic T-lymphocytes.  
 CC -!- DOMAIN: Can be divided into a N-terminal globular domain, a proline-rich P-domain forming an elongated arm-like structure and a C-terminal acidic domain. The P-domain binds one molecule of

CC calcium with high affinity, whereas the acidic C-domain binds  
 CC multiple calcium ions with low affinity (By similarity).  
 CC -|- DOMAIN: The interaction with glycans occurs through a binding site  
 CC in the globular lectin domain (By similarity).  
 CC -|- DOMAIN: The zinc binding sites are localized to the N-domain (By  
 CC similarity).  
 CC -|- DOMAIN: Associates with ERp57 through the tip of the extended arm  
 CC formed by the P-domain (By similarity).  
 CC -|- SIMILARITY: Belongs to the calreticulin family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC -----  
 CC EMBL; X14926; CAA33053.1; -;  
 CC EMBL; M92988; AAA37569.1; -;  
 CC EMBL; AK075605; BAC35852.1; -;  
 CC EMBL; BC003453; AAF03453.1; -;  
 CC PIR; S06763; S06763.  
 CC HSSP; P18418; 1LHN.  
 CC SWISS-2DPAGE; P14211; MOUSE.  
 CC MGD; MGI:88252; Calr.  
 CC GO; GO:0005509; F:calcium ion binding; IDA.  
 CC InterPro; IPR001580; Calret/calnex.  
 CC InterPro; IPR009169; Calreticulin.  
 CC InterPro; IPR009033; Calret calnex P.  
 CC InterPro; IPR008985; ConA like lec\_gl.  
 CC InterPro; IPR000886; ER target\_S.  
 CC Pfam; PF00362; Calreticulin; 1.  
 CC PIRSF; PIRSF002356; Calreticulin; 1.  
 CC PRINTS; PR00626; CALRETICULIN.  
 CC ProDom; PD001866; Calreticulin; 1.  
 CC PROSITE; PS00803; CALRETICULIN 1; 1.  
 CC PROSITE; PS00804; CALRETICULIN 2; 1.  
 CC PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 CC PROSITE; PS00014; ER\_TARGET; 1.  
 CC Calcium-binding; Chaperone; Direct protein sequencing;  
 CC Endoplasmic reticulum; Lectin; Metal-binding; Repeat; Signal; Zinc.  
 CC  
 CC FT SIGNAL 1 17  
 CC FT CHAIN 18 416 Calreticulin.  
 CC FT DOMAIN 18 197 N-domain.  
 CC FT DOMAIN 198 308 P-domain.  
 CC FT DOMAIN 309 416 C-domain.  
 CC FT REPEAT 191 235 4 X approximate repeats.  
 CC FT REPEAT 210 221 1-1.  
 CC FT REPEAT 227 238 1-2.  
 CC FT REPEAT 244 255 1-3.  
 CC FT REPEAT 259 283 1-4.  
 CC FT REPEAT 283 297 3 X approximate repeats.  
 CC FT REPEAT 297 308 2-1.  
 CC FT REPEAT 309 317 2-2.  
 CC FT REPEAT 318 327 2-3.  
 CC FT DOMAIN 351 407 Asp/Glu/Lys-rich.  
 CC FT DISULFID 105 137 By similarity.  
 CC FT SITE 413 416 Prevent secretion from ER.  
 CC SEQUENCE 416 AA; 47994 MW; 24C03B00913408D8 CRC64;

Query Match 95.8%; Score 930; DB 1; Length 416;  
 Best Local Similarity 94.4%; Pred. No. 3.1e-75;  
 Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 EPAYVFEQFLDGDGWTSGWIESKHSDFGKFLVSSGKFGYDGEKDLQTSQDARFVAL 60  
 Db 18 DPAIVFEQFLDGDGWTSGWIESKHSDFGKFLVSSGKFGYDGEKDLQTSQDARFVAL 77  
 QY 61 SASFPFNSKGTLLVQVTFVTHQEQNDICGGGVYKLFNPSLDQTDVHGDSEYNINFGPDIC 120  
 Db 78 SAKFPFNSKGTLLVQVTFVTHQEQNDICGGGVYKLFNPSLDQTDVHGDSEYNINFGPDIC 137

QY 121 GPGTKKVVHVFNYKGNVLINKDIRCKDDEFFTHLYTLIVRPDNTYEVKINDSQQVSGSLE 180  
 Db 138 GPGTKKVVHVFNYKGNVLINKDIRCKDDEFFTHLYTLIVRPDNTYEVKINDSQQVSGSLE 197  
 RESULT 9  
 BAC35852  
 ID BAC35852 PRELIMINARY; PRT; 416 AA.  
 AC BAC35852;  
 DT 14-APR-2004 (TrEMBLrel. 27, Created)  
 DT 14-APR-2004 (TrEMBLrel. 27, Last sequence update)  
 DE Adult male brain cDNA, RIKEN full-length enriched library,  
 DE clone:0710001D13 product:calreticulin, full insert sequence.  
 OS Mus musculus (Mouse)  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RC MEDLINE=22354683; PubMed=12466851;  
 RA The FANTOM Consortium.  
 RA the RIKEN Genome Exploration Research Group Phase I & II Team;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 RT 60,770 full-length cDNAs."  
 RL Nature 420:563-573 (2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RC MEDLINE=21085660; PubMed=11217851;  
 RA RIKEN PANTOM Consortium;  
 RA "Functional annotation of a full-length mouse cDNA collection."  
 RT Nature 409:685-690 (2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RC MEDLINE=99279253; PubMed=10349636;  
 RA Carninci P., Hayashizaki Y.;  
 RA Carninci P., Hayashizaki Y.;  
 RT "High-efficiency full-length cDNA cloning."  
 RL Meth. Enzymol. 303:19-44 (1999).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RC MEDLINE=20493374; PubMed=11042159;  
 RA Carninci P., Shibata Y., Hayatsu M., Sugahara Y., Shibata K., Itoh M.,  
 RA Konno H., Okazaki Y., Muramatsu N., Hayashizaki Y.;  
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to  
 RT prepare full-length cDNA libraries for rapid discovery of new genes."  
 RL Genome Res. 10:1617-1630 (2000).  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RC MEDLINE=20530913; PubMed=11076861;  
 RA Konno H., Akiyama J., Nishi K., Kitsumai T., Tashiro H., Itoh M.,  
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,  
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,  
 RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,  
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,  
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;  
 RT "RIKEN integrated sequence analysis (RISA) system-384-format  
 RT sequencing pipeline with 384 multicapillary sequencer."  
 RL Genome Res. 10:1757-1771 (2000).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Brain;  
 RC Aichi J., Aizawa K., Akahira S., Akimura T., Aono H., Arai A.,  
 RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,  
 RA Hanaoka T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,  
 RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,  
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,  
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,



```

RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamanura T., Yamanaka I., Yasunishi A.,
RA Yoshida K., Yoshino M., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK075605; BAC35852.1; - - - 24C03B00913408D8 CRC64;
SQ SEQUENCE 416 AA; 47994 MW; 24C03B00913408D8 CRC64;

Query Match 95.8%; Score 930; DB 2; Length 416;
Best Local Similarity 94.4%; Pred. No. 3.1e-75;
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPAVYFKQFLDGDGWTNRWIESKHKSDRGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60
DB 18 DPAIYFKQFLDGDGWTNRWIESKHKSDRGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASFEPFNSKGTQVLVQVFTVKHEQNDICGGGVYKLFPSNLDQDMDHGDSEYNIMFGPDIC 120
DB 78 SAKFEPFNSKGTQVLVQVFTVKHEQNDICGGGVYKLFPSNLDQDMDHGDSEYNIMFGPDIC 137

QY 121 GPQTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLVLRPNTYEVKIDNSQVSGSLE 180
DB 138 GPQTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLVLRPNTYEVKIDNSQVSGSLE 197

RESULT 10
CRIT_BOVIN STANDARD; PRT; 417 AA.
AC P52193; Q8S053;
DT 01-OCT-1996 (Rel. 34, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Calreticulin, brain isoform 1 precursor (CRP55) (Calregulin) (HACBP).
GN Name=CALR; Synonyms=CRT;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RC TISSUE=Brain;
RA Hossain M.A., Takuwa K., Minakata H., Nakajima T.;
RT "Bovine brain calreticulin.";
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE OF 18-417.
RC TISSUE=Brain;
RX MEDLINE=9418174; PubMed=8135753;
RA Matsuoka K., Seta K., Yamakawa Y., Okuyama T., Shinoda T., Isobe T.;
RT "Covalent structure of bovine brain calreticulin.";
RL Biochem. J. 298:435-442(1994).
CC 1- FUNCTION: Molecular calcium binding chaperone promoting folding,
CC oligomeric assembly, and quality control in the ER via the
CC calreticulin/calnexin cycle. This lectin interacts transiently
CC with almost all of the monoglucosylated glycoproteins that are
CC synthesized in the ER. Interacts with the DNA-binding domain of
CC NR3C1 and mediates its nuclear export (By similarity).
CC 1- SUBUNIT: Monomer. Interacts with Grp58 (Erp57) and with NR3C1 (By
CC similarity).
CC 1- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC 1- DOMAIN: Can be divided into a N-terminal globular domain, a
CC proline-rich P-domain forming an elongated arm-like structure and
CC a C-terminal acidic domain. The P-domain binds one molecule of
CC calcium with high affinity, whereas the acidic C-domain binds
CC multiple calcium ions with low affinity (By similarity).
CC 1- DOMAIN: The interaction with glycans occurs through a binding site
CC in the globular lectin domain (By similarity).
CC 1- DOMAIN: The zinc binding sites are localized to the N-domain (By
CC similarity).
CC 1- DOMAIN: Associates with Erp57 through the tip of the extended arm
CC formed by the P-domain (By similarity).

```

```

CC 1- SIMILARITY: Belongs to the calreticulin family.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/ebis/
CC or send an email to license@ebi.ac.uk).
CC EMBL; AB067687; BAB86913.1; - - -
DR HSSP; P18418; 1K91.
DR InterPro; IPR001580; Calret/calnex.
DR InterPro; IPR009169; Calreticulin.
DR InterPro; IPR009033; Calret calnex P.
DR InterPro; IPR008985; ConA like lec_g1.
DR InterPro; IPR008886; ER_target_S.
DR Pfam; PF00262; Calreticulin; 1.
DR PIRSF; PIRSF002356; Calreticulin; 1.
DR PRINTS; PC00626; CALRETICULIN.
DR ProDom; PD001866; Calreticulin; 1.
DR PROSITE; PS00803; CALRETICULIN_1; 1.
DR PROSITE; PS00804; CALRETICULIN_2; 1.
DR PROSITE; PS00805; CALRETICULIN_REPEAT; 3.
DR PROSITE; PS00014; ER_TARGET; 1.
KW Calcium-binding; Chaperone; Direct protein sequencing;
KW Endoplasmic reticulum; Glycoprotein; Lectin; Metal-binding; Repeat;
KW Signal; Zinc.
FT SIGNAL 1 17 Calreticulin, brain isoform 1.
FT CHAIN 18 417 N-domain.
FT DOMAIN 18 197 P-domain.
FT DOMAIN 198 308 C-domain.
FT DOMAIN 191 255 4 X approximate repeats.
FT REPEAT 191 202 1-1.
FT REPEAT 202 221 1-2.
FT REPEAT 221 238 1-3.
FT REPEAT 238 255 1-4.
FT REPEAT 255 297 3 X approximate repeats.
FT DOMAIN 259 269 2-1.
FT REPEAT 269 273 2-2.
FT REPEAT 273 283 2-3.
FT REPEAT 283 297 2-3.
FT DOMAIN 351 407 Asp/Glu/Lys-rich.
FT DISULFID 179 163 N-linked (GlcNAc...).
FT CARBOHYD 179 179 Prevent secretion from ER (Potential).
FT SITE 414 417
SQ SEQUENCE 417 AA; 48038 MW; 7BF812C7B54178E9 CRC64;

Query Match 95.8%; Score 930; DB 1; Length 417;
Best Local Similarity 95.0%; Pred. No. 3.1e-75;
Matches 171; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPAVYFKQFLDGDGWTNRWIESKHKSDRGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60
DB 18 DPAIYFKQFLDGDGWTNRWIESKHKSDRGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASFEPFNSKGTQVLVQVFTVKHEQNDICGGGVYKLFPSNLDQDMDHGDSEYNIMFGPDIC 120
DB 78 SAKFEPFNSKGTQVLVQVFTVKHEQNDICGGGVYKLFPSNLDQDMDHGDSEYNIMFGPDIC 137

QY 121 GPQTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLVLRPNTYEVKIDNSQVSGSLE 180
DB 138 GPQTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLVLRPNTYEVKIDNSQVSGSLE 197

RESULT 11
Q6EE32 PRELIMINARY; PRT; 404 AA.
ID Q6EE32
AC Q6EE32; (TrEMBLrel. 28, Created)
DT 01-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 01-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Calreticulin (Fragment).

```

SEQUENCE FROM N.A.  
 TISSUE=Embryo;  
 MEDLINE=22341132; PubMed=12454917;  
 Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 Richardson P.;  
 "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 Dev. Dyn. 225:384-391(2002)."  
 [3]  
 SEQUENCE FROM N.A.  
 TISSUE=Embryo;  
 Klein S., Strausberg R.;  
 Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.  
 CC -I- SIMILARITY: Belongs to the calreticulin family.  
 DR EMBL; BC046699; AAH46699.1; -;  
 HSP; P18418; IK91.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR GO; GO:0003754; F:chaperone activity; IEA.  
 DR InterPro; IPRO01580; Calret/calnex.  
 DR InterPro; IPRO09169; Calreticulin.  
 DR InterPro; IPRO09033; Calret.calnex.p.  
 DR InterPro; IPRO08985; ConA like lec\_g1.  
 DR InterPro; IPRO00886; ER\_target\_S\_g1.  
 DR PIRSF; PIRSF002356; calreticulin; 1.  
 DR PRINTS; PR00625; CALRETICULIN.  
 DR ProDom; PD001866; Calret/calnex; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 3.  
 DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
 DR Chapsone; Endoplasmic reticulum.  
 KW SEQUENCE 413 AA; 48521 MW; DE002F8F0523772B CRC64;  
 Query Match 86.8%; Score 843; DB 2; Length 413;  
 Best Local Similarity 84.4%; Pred. No. 2,1e-67;  
 Matches 152; Conservative 14; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 EBAVYFKEGFLDGDGWTSRWISKHSKSPGKVLSSGKFGYDEEKDKGLQTSQDARFYAL 60  
 DB 19 EBAVYFKEEFGDGDGWTQWWSKHSKSYGKFKLSAGKFGYSEKDKGLQTSQDARFYAM 78  
 QY 61 SASFPFNSKGTGLVYQFVTKHEQNIQCGGVYKVLFPNSLDOTDMHGDSYNNIMFGPDIC 120  
 DB 79 SRFPFNSKGTGLVYQFSVKHEQNIQCGGVYKVLFPADLEGTENHGSEYNNIMFGPDIC 138  
 QY 121 GPGTKKVVHVFNYKGNVLINKDKDDFFHLYTLVIRPDNTYEVKIDNSQVSGSLE 180  
 DB 139 GPFTKVVHVFQYKKKNLQINKDKDDSFTHLYTLVIRPDNTYEVKIDNSKVSGSLE 198  
 RESULT 13  
 Q7ZYX3  
 ID Q7ZYX3 PRELIMINARY; PRT; 418 AA.  
 AC Q7ZYX3;  
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
 DT 01-VAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DE Crc-prov protein.  
 DE OS Xenopus laevis (African clawed frog).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;  
 OC Xenopodinae; Xenopus.  
 OX NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22388257; PubMed=12477932;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heich F.,  
 RA Diachenko I., Varsina K., Farver A.A., Rubin G.M., Heath  
 RA

01-MAR-2004 (TrEMBLrel. 26, Last annotation update)  
 DT Calreticulin precursor (Fragment).  
 DE Xenopus laevis (African clawed frog).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 CC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8355;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=CNS;  
 RA Treves S.; Zorzato F.; Pozzan T.;  
 RT "Identification of calreticulin isoform in the CNS."  
 RL Biochem. J. 0:0-0(0).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=CNS;  
 RA Treves S.;  
 RL Submitted (JUL-1992) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the calreticulin family.  
 DR EMBL; X67597; CAA47866.1; -;  
 DR PIR; S29129; S29129.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR InterPro; IPR001580; Calret/calnex.  
 DR InterPro; IPR003033; Calret/calnex.P.  
 DR InterPro; IPR003985; ConA-like lec\_gl.  
 DR InterPro; IPR000886; ER target\_S.  
 DR Pfam; PF00262; Calreticulin; 1.  
 DR PRINTS; PRO0626; CALRETICULIN.  
 DR ProDom; PDC01866; Calret/calnex; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 2.  
 DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
 DR Chaperone; Endoplasmic reticulum; Signal.  
 KW NON\_TER  
 FT SIGNAL  
 FT CHAIN  
 SQ SEQUENCE 411 AA; 48344 MW; 831DA66E0CEBEFA CRC64;  
 Query Match 86.0%; Score 835; DB 2; Length 411;  
 Best Local Similarity 83.9%; Pred. No. 1.1e-66;  
 Matches 151; Conservative 15; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 EPVYFKEQFLDGGWTSRWIESKHGDFGVLSGKFGYDSEKDKGLQTSQDARFYAL 60  
 Db 13 EPVYFKEEFTDGGWTSRWIESKHGDFGVLSGKFGYDSEKDKGLQTSQDARFYAM 72  
 QY 61 SASPEFFSNKQTLVQVTFVKHEQNIIDCGGGYVKLPFNSLDQTMHGDSEYNIMFGPDIC 120  
 Db 73 SSRPDSFSNKDQTLVQVFSVKHEQNIIDCGGGYVKLPFPAALQETEMHSESEYNIMFGPDIC 132  
 QY 121 GPCTKKVHVFNYKGNVINKIDCKDDETHLYTLVLPDNTVEVKIDNSOVESGSLE 180  
 Db 133 GPCTKKVHVFQYKKKQNLQINKIDCKDDETHLYTLVLPDNTVEVKIDNSOVESGSLE 192  
 RESULT 15  
 Q6NVIT6 PRELIMINARY; PRT; 416 AA.  
 ID Q6NVIT6  
 AC Q6NVIT6;  
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)  
 DE Hypothetical protein MGC69541.  
 GN Name=MGC69541;  
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;  
 CC Xenopodinae; Xenopus.  
 NCBI\_TaxID=8364;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mollany S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,  
 RA Jones S.J., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences."  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RX MEDLINE=22341132; PubMed=12454917;  
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,  
 RA Richardson P.;  
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus  
 RT initiative."  
 RL Dev. Dyn. 225:384-391(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Embryo;  
 RA Klein S., Strausberg R.;  
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Belongs to the calreticulin family.  
 DR EMBL; BC040668; AA44068.1; -;  
 DR HSSP; P18418; IK91.  
 DR GO; GO:0005783; C:endoplasmic reticulum; IEA.  
 DR GO; GO:0005514; F:calcium ion storage activity; IEA.  
 DR GO; GO:0003754; F:chaperone activity; IEA.  
 DR InterPro; IPR001580; Calret/calnex.  
 DR InterPro; IPR009169; Calreticulin.  
 DR InterPro; IPR009033; Calret/calnex.P.  
 DR InterPro; IPR008985; ConA-like lec\_gl.  
 DR InterPro; IPR000886; ER target\_S.  
 DR PIRSF; PIRSF002356; Calreticulin; 1.  
 DR PRINTS; PRO0626; CALRETICULIN.  
 DR ProDom; PDC01866; Calret/calnex; 1.  
 DR PROSITE; PS00803; CALRETICULIN\_1; 1.  
 DR PROSITE; PS00804; CALRETICULIN\_2; 1.  
 DR PROSITE; PS00805; CALRETICULIN\_REPEAT; 2.  
 DR PROSITE; PS00014; ER\_TARGET; UNKNOWN\_1.  
 KW Chaperone; Endoplasmic reticulum.  
 SQ SEQUENCE 418 AA; 49028 MW; 731C1C9AA03F9A53 CRC64;  
 Query Match 86.8%; Score 843; DB 2; Length 418;  
 Best Local Similarity 84.4%; Pred. No. 2.1e-67;  
 Matches 152; Conservative 15; Mismatches 13; Indels 0; Gaps 0;  
 QY 1 EPVYFKEQFLDGGWTSRWIESKHGDFGVLSGKFGYDSEKDKGLQTSQDARFYAL 60  
 Db 19 EPVYFKEEFTDGGWTSRWIESKHGDFGVLSGKFGYDSEKDKGLQTSQDARFYAM 78  
 QY 61 SASPEFFSNKQTLVQVTFVKHEQNIIDCGGGYVKLPFNSLDQTMHGDSEYNIMFGPDIC 120  
 Db 79 SSRPDSFSNKDQTLVQVFSVKHEQNIIDCGGGYVKLPFPAALQETEMHSESEYNIMFGPDIC 138  
 QY 121 GPCTKKVHVFNYKGNVINKIDCKDDETHLYTLVLPDNTVEVKIDNSOVESGSLE 180  
 Db 139 GPCTKKVHVFQYKKKQNLQINKIDCKDDETHLYTLVLPDNTVEVKIDNSOVESGSLE 198  
 RESULT 14  
 Q91710 PRELIMINARY; PRT; 411 AA.  
 ID Q91710  
 AC Q91710;  
 DT 01-JAN-1998 (TrEMBLrel. 05, Created)  
 DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)

Search completed: October 15, 2004, 13:12:33  
Job time : 68 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2004, 13:11:17 ; Search time 15 Seconds  
(without alignments)  
1154.601 Million cell updates/sec

Title: US-09-807-148-4  
Perfect score: 971  
Sequence: 1 EFAVYKEQFLDGDGWTNRW.....PDNTVEVKIDNSQVSGSL 180

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:.\*  
1: PIR1:.\*  
2: PIR2:.\*  
3: PIR3:.\*  
4: PIR4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	971	100.0	417	1 A37047	calreticulin precu
2	933	96.1	416	2 JH0819	calreticulin precu
3	931	95.9	418	1 A34154	calreticulin precu
4	930	95.8	400	2 S43176	calreticulin, brai
5	930	95.8	416	1 S06763	calreticulin precu
6	835	86.0	411	2 S29129	calreticulin precu
7	804	82.8	405	1 JH0795	calreticulin precu
8	797	82.1	384	2 S29130	calreticulin (clon
9	796	82.0	419	2 S71343	calreticulin precu
10	699	72.0	406	2 A56637	calreticulin homol
11	619.5	63.8	395	2 S25851	calreticulin precu
12	542.5	55.9	336	2 A32507	41K larval antigen
13	539	55.5	412	2 T05703	calreticulin - bar
14	539	55.5	415	2 T05705	calreticulin - bar
15	536	55.2	421	2 S58170	calreticulin precu
16	531	54.7	416	2 T14554	calreticulin - bee
17	527	54.3	415	2 T10172	calreticulin - cas
18	522.5	53.8	421	2 S36799	calreticulin precu
19	522	53.8	416	2 T16968	calreticulin call
20	517	53.2	389	2 T03691	calreticulin - com
21	512	52.7	425	2 C96605	calreticulin (Crt1
22	510	52.5	444	2 H86224	hypothetical prote
23	496	51.1	393	1 A48573	calreticulin autoa
24	347	35.7	422	2 T07841	probable calreticu
25	293	30.2	591	2 B54344	calnexin precursor
26	290	29.9	591	2 C54354	calnexin precursor
27	290	29.9	592	2 F53260	calnexin - human
28	290	29.9	593	1 A37273	calnexin precursor
29	288	29.7	592	2 A46673	calnexin precursor

30	270	27.8	611	2 A53418	calnexin precursor
31	262	27.0	622	2 S71342	calnexin precursor
32	253.5	26.1	560	2 S56142	calcium-binding pr
33	252	26.0	611	2 A54066	calnexin-t - mouse
34	248	25.5	582	2 A46637	calnexin homolog S
35	244	25.1	619	2 S40938	calnexin-like prot
36	238.5	24.6	530	2 JN0597	calnexin-like prot
37	232	23.9	546	2 T06415	calnexin - soybean
38	215.5	22.2	540	2 T10892	probable calnexin
39	200.5	20.6	532	2 T49873	calnexin homolog -
40	157	16.2	29	2 E33208	calreticulin, uter
41	145	14.9	29	2 C33208	calreticulin, slow
42	144	14.8	29	2 D33208	calreticulin, brai
43	133.5	13.7	502	2 S29347	calnexin homolog Y
44	129	13.3	297	2 S70552	calnexin homolog C
45	124.5	12.8	428	2 T03251	calnexin - maize (

ALIGNMENTS

RESULT 1

A37047  
calreticulin precursor - human  
N:Alternate names: 52K ribonucleoprotein autoantigen Ro/SS-A; 60K integrin-binding prote  
C:Species: Homo sapiens (man)  
C:Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text change 09-Jul-2004  
C:Accession: A42330; A37047; A46452; A28812; PH1525; A40346; S11475; T45075  
R:McCaulliffe, D.P.; Yang, Y.S.; Wilson, J.; Sontheimer, R.D.; Capra, J.D.  
J. Biol. Chem. 267, 2557-2562, 1992  
A:Title: The 5'-flanking region of the human calreticulin gene shares homology with the  
A:Reference number: A42330; MUID:92129342; PMID:1733953  
A:Accession: A42330  
A:Molecule type: DNA  
A:Residues: 1-417 <MC2>  
A:Cross-references: UNIPROT:P27797  
A:Note: sequences extracted from NCBI backbone (NCBIN:78537, NCBIP:78536)  
R:McCaulliffe, D.P.; Lux, F.A.; Lieu, T.S.; Sanz, I.; Hanke, J.; Newkirk, M.M.; Bachinski  
J. Clin. Invest. 85, 1379-1391, 1990  
A:Title: Molecular cloning, expression, and chromosome 19 localization of a human Ro/SS  
A:Reference number: A37047; MUID:90237213; PMID:2332496  
A:Accession: A37047  
A:Molecule type: mRNA  
A:Residues: 1-417 <MC2>  
A:Cross-references: GB:M32294; NID:9337486; PIDN:AAA36582.1; PID:G337487  
R:Roach, L.A.; Haselby, J.A.; Meilof, J.F.; Smeenk, R.J.; Unnasch, T.R.; Greene, B.M.,  
J. Immunol. 147, 3031-3039, 1991  
A:Title: Characterization of the autoantigen calreticulin.  
A:Reference number: A46452; MUID:92013129; PMID:1919005  
A:Accession: A46452  
A:Molecule type: mRNA  
A:Residues: 1-417 <R0K>  
A:Cross-references: GB:M84739; NID:9179881; PIDN:AAA51916.1; PID:G179882  
A:Note: sequence extracted from NCBI backbone (NCBIN:60749, NCBIP:60750)  
R:Lieu, T.S.; Newkirk, M.M.; Capra, J.D.; Sontheimer, R.D.  
J. Clin. Invest. 82, 96-101, 1988  
A:Title: Molecular characterization of human Ro/SS-A antigen. Amino terminal sequence o  
A:Reference number: A28812; MUID:89273610; PMID:3260607  
A:Accession: A28812  
A:Molecule type: protein  
A:Residues: 18-41 <LIE>  
A:Note: 18-Ala was also found  
R:Dupuis, M.; Schaefer, E.; Krause, K.H.; Tschopp, J.  
J. Exp. Med. 177, 1-7, 1993  
A:Title: The calcium-binding protein calreticulin is a major constituent of lytic granu  
A:Reference number: PH1525; MUID:93115648; PMID:8418194  
A:Accession: PH1525  
A:Molecule type: protein  
A:Residues: 18-27 <DUP>  
A:Experimental source: LAK cell  
R:Rojiani, M.V.; Finlay, B.B.; Gray, V.; Dedhar, S.  
Biochemistry 30, 9859-9866, 1991

A;Title: In vitro interaction of a polypeptide homologous to human Ro/SS-A antigen (calc  
A;Reference number: A40346; MUID:92002034; PMID:1911778  
A;Accession: A40346  
A;Molecule type: protein  
A;Residues: 18-34; <R>  
R;Krause, K.H.; Simmerman, H.K.B.; Jones, L.R.; Campbell, K.P.  
Biochem. J. 270, 545-548, 1990  
A;Title: Sequence similarity of calreticulin with a Ca(2+)-binding protein that co-purif  
A;Reference number: S11475; MUID:90380058; PMID:2400400  
A;Accession: S11475  
A;Molecule type: protein  
A;Residues: 18-32 <R>  
R;Immerdin, J.; McCready, P.; Stilwagen, S.; Ramirez, M.; Carrano, A.  
submitted to the EMBL Data Library, November 1996  
A;Description: Characterization by genomic sequence analysis of a gene-rich 111 kb regio  
A;Reference number: Z22906  
A;Accession: T45075  
A;Status: preliminary; translated from GB/EMBL/DDBJ  
A;Molecule type: DNA  
A;Residues: 1-417 <L>  
A;Cross-references: EMBL:AD000092; PIDN:AAB51176.1  
A;Experimental source: Cell line SH2-B; fibroblast  
C;Comment: Autoantibodies specific for this protein are found in Sjogren's syndrome and  
C;Genetics:  
A;Gene: GDB:CALR  
A;Cross-references: GDB:125179; OMIM:109091  
A;Map position: 19p13.3-19p13.2  
A;Intons: 31/1; 65/1; 133/1; 164/3; 234/3; 272/3; 320/3; 351/3  
A;Note: CRT  
C;Superfamily: calreticulin  
C;Keywords: calcium binding; integrin binding  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-417/Product: calreticulin #status predicted <MAT>  
F;414-417/Region: endoplasmic reticulum retention signal  
Query Match 100.0%; Score 971; DB 1; Length 417;  
Best Local Similarity 100.0%; Pred. No. 4.3e-81;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EBAVYFKEQFLDGDGWTSEWIESKHSDFGKVLSSGFYGDDEKKGLOTSDQARFYAL 60  
DB 18 EBAVYFKEQFLDGDGWTSEWIESKHSDFGKVLSSGFYGDDEKKGLOTSDQARFYAL 77  
QY 61 SASFPFNSKNGQTLVQVTFVKEQNI DCGGVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
DB 78 SASFPFNSKNGQTLVQVTFVKEQNI DCGGVKLPNSLDQTMHGDSEYNIMFGPDIC 137  
QY 121 GPGTKKHVHFYNGKXNVLINKDIRCKDDEFTLTVLRPDNTVEVKIDNSQVESGSL 180  
DB 138 GPGTKKHVHFYNGKXNVLINKDIRCKDDEFTLTVLRPDNTVEVKIDNSQVESGSL 197  
RESULT 2  
JH0819  
calreticulin precursor - rat  
N;Alternate names: calcium-binding protein 3  
C;Species: Rattus norvegicus (Norway rat)  
C;Date: 30-Sep-1993 #sequence revision 20-Aug-1994 #text change 09-Jul-2004  
C;Accession: JH0819; A49176; S11205; PC1109; S45036; S04867; S39372; A34473; S13045  
R;Nakamura, M.; Moriya, M.; Baba, T.; Michikawa, Y.; Imanobe, T.; Arai, K.; Okinaga, S.  
Exp. Cell Res. 205, 101-110, 1993  
A;Title: An endoplasmic reticulum protein, calreticulin, is transported into the acrosom  
A;Reference number: A49176; MUID:93202172; PMID:8453984  
A;Accession: JH0819  
A;Molecule type: mRNA  
A;Residues: 1-416 <NA>  
A;Cross-references: UNIPROT:P18418; GB:D78308; NID:g1089798; PIDN:BA11345.1; PID:g18455  
A;Accession: A49176  
A;Status: preliminary  
A;Molecule type: mRNA; protein  
A;Residues: 1-416 <NA>  
A;Cross-references: GB:D78308; NID:g1089798; PIDN:BA11345.1; PID:g1845572  
A;Experimental source: Sprague-Dawley, spermatogenic cells

A;Note: sequence extracted from NCBI backbone (NCBIN:127639, NCBI:127643)  
R;Murthy, K.K.; Banville, D.; Srikant, C.B.; Carrier, F.; Holmes, C.; Bell, A.; Patel, Y.  
Nucleic Acids Res. 18, 4933, 1990  
A;Title: Structural homology between the rat calreticulin gene product and the Onchocerc  
A;Reference number: S11205; MUID:90370496; PMID:2995661  
A;Accession: S11205  
A;Molecule type: mRNA  
A;Residues: 1-416 <NA>  
A;Cross-references: EMBL:X53363; NID:G55854; PIDN:CAA37446.1; PID:G55855  
R;Nakamura, M.; Michikawa, Y.; Baba, T.; Okinaga, S.; Arai, K.  
Biochem. Biophys. Res. Commun. 186, 668-673, 1992  
A;Title: Calreticulin is present in the acrosome of spermatozoa of rat testis.  
A;Reference number: PC1109; MUID:92360010; PMID:1497655  
A;Accession: PC1109  
A;Molecule type: protein  
A;Residues: 18-32 <NA>  
A;Experimental source: testis, strain Sprague-Dawley  
R;Seennichsen, B.; Fuellekrug, J.; van Nguyen, P.; Diekmann, W.; Robinson, D.G.; Mieskes  
submitted to the EMBL Data Library, May 1994  
A;Description: Retention and retrieval: both mechanisms cooperate to maintain calreticul  
A;Reference number: S45036  
A;Accession: S45036  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-416 <SOE>  
A;Cross-references: EMBL:X79327; NID:G488840; PIDN:CAAS5890.1; PID:G488841  
R;Lone, Y.C.; Bailly, A.; Latruffe, N.  
submitted to the EMBL Data Library, December 1988  
A;Reference number: S04867  
A;Accession: S04867  
A;Molecule type: mRNA  
A;Residues: R', 270-358, 'AAG' <LON>  
A;Cross-references: EMBL:X13702; NID:G56055; PIDN:CAA31387.1; PID:G930260  
A;Note: the authors designated the protein as D-beta-hydroxybutyrate dehydrogenase  
R;Yokoi, T.; Nagayama, S.; Kajiwar, R.; Kawaguchi, Y.; Horiuchi, R.; Kamataki, T.  
Biochim. Biophys. Acta 1158, 339-344, 1993  
A;Title: Identification of protein disulfide isomerase and calreticulin as autoimmune an  
A;Reference number: S39371; MUID:94072621; PMID:8251535  
A;Accession: S39372  
A;Molecule type: protein  
A;Residues: 18-23, 'X', 25-32 <YOK>  
R;Van, P.N.; Peter, F.; Soelling, H.D.  
J. Biol. Chem. 264, 17494-17501, 1989  
A;Title: Four intracisternal calcium-binding glycoproteins from rat liver microsomes wit  
itive calcium sequestering rat liver vesicles.  
A;Reference number: A34473; MUID:90008920; PMID:2793869  
A;Accession: A34473  
A;Status: preliminary  
A;Molecule type: protein  
A;Residues: 18-36 <VAN>  
R;Treves, S.; de Mattei, M.; Lanfredi, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld  
Biochem. J. 271, 473-480, 1990  
A;Title: Calreticulin is a candidate for a calsequestrin-like function in Ca(2+)-storage  
A;Reference number: S13045; MUID:91054414; PMID:2241926  
A;Accession: S13045  
A;Molecule type: protein  
A;Residues: 18-29 <TRE>  
C;Superfamily: calreticulin  
C;Keywords: calcium binding; glycoprotein  
F;1-17/Domain: signal sequence #status predicted <SIG>  
F;18-416/Product: calreticulin #status experimental <MAT>  
F;204-212/Region: nuclear location signal  
F;413-416/Region: endoplasmic reticulum retention signal  
F;344/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 96.1%; Score 933; DB 2; Length 416;  
Best Local Similarity 94.4%; Pred. No. 1.3e-77;  
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;  
QY 1 EBAVYFKEQFLDGDGWTSEWIESKHSDFGKVLSSGFYGDDEKKGLOTSDQARFYAL 60  
DB 18 DPAIFKQFLDGDGWTSEWIESKHSDFGKVLSSGFYGDDEKKGLOTSDQARFYAL 77

QY 61 SASPEPNSKQTLVQFTVKHEQNIIDCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
 DB 78 SARPEPNSKQTLVQFTVKHEQNIIDCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 137  
 QY 121 GPQTKKVHVFYNYKGNVLINKIDCKDDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 180  
 DB 138 GPQTKKVHVFYNYKGNVLINKIDCKDDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 197

RESULT 3  
 A34154  
 A: calreticulin precursor, skeletal muscle - rabbit  
 C: Species: Oryctolagus cuniculus (domestic rabbit)  
 C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C: Accession: A34154; S13047  
 R: Plietel, L.; Burns, K.; MacLennan, D.H.; Reithmeier, R.A.F.; Michalak, M.  
 J. Biol. Chem. 264, 21522-21528, 1989  
 A: Title: Molecular cloning of the high affinity calcium-binding protein (calreticulin) c  
 A: Reference number: A34154; MUID:900943320; PMID:2600080  
 A: Accession: A34154  
 A: Status: preliminary  
 A: Molecule type: mRNA  
 A: Residues: 1-418 <FLI>  
 A: Cross-references: UNIPROT:P15253; GB:J05138; NID:G164858; PIDN:AAA31188.1; PID:G164858  
 R: Treves, S.; de Mattei, M.; Lanfretti, M.; Villa, A.; Green, N.M.; MacLennan, D.H.; Meld  
 Biochem. J. 271, 473-480, 1990  
 A: Title: Calreticulin is a candidate for a caldesmon-like function in Ca(2+)-storage  
 A: Reference number: S13045; MUID:91054414; PMID:2241926  
 A: Accession: S13047  
 A: Molecule type: protein  
 A: Residues: 19-32 <PRE>  
 C: Superfamily: calreticulin  
 C: Keywords: skeletal muscle  
 F: 1-17/Domain: signal sequence #status predicted <SIG>  
 F: 415-418/Region: endoplasmic reticulum retention signal

Query Match 95.9%; Score 931; DB 1; Length 418;  
 Best Local Similarity 95.8%; Pred. No. 1.9e-77;  
 Matches 172; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 1 EPAYVFKQFLDGGWTSRWIESKHSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
 DB 18 EPVYVFKQFLDGGWTSRWIESKHSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77  
 QY 61 SASPEPNSKQTLVQFTVKHEQNIIDCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
 DB 78 SARPEPNSKQTLVQFTVKHEQNIIDCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 137  
 QY 121 GPQTKKVHVFYNYKGNVLINKIDCKDDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 180  
 DB 138 GPQTKKVHVFYNYKGNVLINKIDCKDDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 197

RESULT 4  
 S43376  
 A: calreticulin, brain isoform 1 - bovine  
 C: Species: Bos primigenius taurus (cattle)  
 C: Date: 20-Oct-1994 #sequence\_revision 23-Mar-1995 #text\_change 07-May-1999  
 C: Accession: S43376; S36801  
 R: Matsuo, K.; Seta, K.; Yamakawa, Y.; Okuyama, T.; Shinoda, T.; Isobe, T.  
 Biochem. J. 298, 435-442, 1994  
 A: Title: Covalent structure of bovine brain calreticulin.  
 A: Reference number: S43376; MUID:94183174; PMID:8135753  
 A: Accession: S43376  
 A: Molecule type: protein  
 A: Residues: 1-400 <Mar>  
 A: Experimental source: brain  
 R: Liu, N.; Fine, R.E.; Johnson, R.J.  
 Biochim. Biophys. Acta 1202, 70-76, 1993  
 A: Title: Comparison of cDNAs from bovine brain coding for two isoforms of calreticulin.  
 A: Reference number: S36799; MUID:93385184; PMID:8373827  
 A: Accession: S36801  
 A: Status: nucleic acid sequence not shown

A: Molecule type: mRNA  
 A: Residues: 45-63, 'E', 65-83 <LIU>  
 A: Experimental source: brain, clone 8.1  
 C: Superfamily: calreticulin  
 C: Keywords: calcium binding; glycoprotein  
 F: 397-400/Region: endoplasmic reticulum retention signal  
 F: 120-146/Disulfide bonds: #status experimental  
 F: 162/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 95.8%; Score 930; DB 2; Length 400;  
 Best Local Similarity 95.0%; Pred. No. 2.3e-77;  
 Matches 171; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 EPAYVFKQFLDGGWTSRWIESKHSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
 DB 1 DPTVYFKQFLDGGWTSRWIESKHSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
 QY 61 SASPEPNSKQTLVQFTVKHEQNIIDCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
 DB 61 SARPEPNSKQTLVQFTVKHEQNIIDCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
 QY 121 GPQTKKVHVFYNYKGNVLINKIDCKDDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 180  
 DB 121 GPQTKKVHVFYNYKGNVLINKIDCKDDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE 180

RESULT 5  
 S06763  
 A: calreticulin precursor - mouse  
 N: Alternate names: 55K calcium-binding reticuloplasmin; calregulin  
 C: Species: Mus musculus (house mouse)  
 C: Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C: Accession: S06763; JCI1444; FCI1233; A57498  
 R: Smith, M.J.; Koch, G.L.E.  
 EMBO J. 8, 3581-3586, 1989  
 A: Title: Multiple zones in the sequence of calreticulin (CRP55, calregulin, HACBP), a m  
 A: Reference number: S06763; MUID:90059955; PMID:2583110  
 A: Accession: S06763  
 A: Molecule type: DNA  
 A: Residues: 1-416 <SMI>  
 A: Cross-references: UNIPROT:P14211; EMBL:X14926; NID:G50567; PIDN:CAA33053.1; PID:G50567  
 R: Mazarelli, R.A.; Gold, P.; Cunningham, M.; Green, M.  
 Gene 120, 217-225, 1992  
 A: Title: Determination of the sequence of an expressible cDNA clone encoding BRP60/calr  
 A: Reference number: JCI1444; MUID:93013037; PMID:1398135  
 A: Accession: JCI1444  
 A: Molecule type: mRNA  
 A: Residues: 1-416 <MA2>  
 A: Cross-references: GB:M92988; NID:G193084; PIDN:AAA37569.1; PID:G193085  
 A: Accession: PC1233  
 A: Molecule type: protein  
 A: Residues: 18-41 <MA2>  
 R: White, T.K.; Zhu, Q.; Tanzer, M.L.  
 J. Biol. Chem. 270, 15926-15929, 1995  
 A: Title: Cell surface calreticulin is a putative mannose lectin which triggers mouse  
 A: Reference number: A57498; MUID:95332280; PMID:7608143  
 A: Accession: A57498  
 A: Status: preliminary  
 A: Molecule type: protein  
 A: Residues: 74-80; 142-151; 186-193 <WHI>  
 C: Superfamily: calreticulin  
 C: Keywords: calcium binding  
 F: 1-17/Domain: signal sequence #status predicted <SIG>  
 F: 18-416/Product: calregulin #status experimental <MAT>  
 F: 413-416/Region: endoplasmic reticulum retention signal

Query Match 95.8%; Score 930; DB 1; Length 416;  
 Best Local Similarity 94.4%; Pred. No. 2.4e-77;  
 Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPAYVFKQFLDGGWTSRWIESKHSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
 DB 18 DPALVFKQFLDGGWTSRWIESKHSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEPSNKGQTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDIC 120  
 Db 78 SAKFPEPSNKGQTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDIC 137  
 QY 121 GPCTKKVHVIFNYKGNVLINKIRCKDDEFTHLYTLIVRPNTVEVKIDNSQVSGSLE 180  
 Db 138 GPCTKKVHVIFNYKGNVLINKIRCKDDEFTHLYTLIVRPNTVEVKIDNSQVSGSLE 197  
 RESULT 6  
 S29129  
 calreticulin precursor (clone 3) - African clawed frog (fragment)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: S29129  
 R:Trevies, S.; Zorzato, F.; Pozzan, T.  
 Biochem. J. 287, 579-581, 1992  
 A:Title: Identification of calreticulin isoforms in the central nervous system.  
 A:Reference number: S29129; MUID:93074997; PMID:1445218  
 A:Accession: S29129  
 A:Molecule type: mRNA  
 A:Residues: 1-411 <TR>  
 A:Cross-references: UNIPROT:Q91710; EMBL:X67597; NID:G64608; PIDN:CAA47866.1; PID:G64609  
 C:Superfamily: calreticulin  
 C:Keywords: glycoprotein  
 F:1-12/Domain: signal sequence (fragment) #status predicted <SIG>  
 F:13-411/Product: calreticulin #status predicted <MAT>  
 F:408-411/Region: endoplasmic reticulum retention signal  
 F:339/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 86.0%; Score 835; DB 2; Length 411;  
 Best Local Similarity 83.9%; Pred. No. 1.1e-68;  
 Matches 151; Conservative 15; Mismatches 14; Indels 0; Gaps 0;  
 QY 1 EPAVYFKEQFLDGDGWTSMWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60  
 Db 13 EPAVYFKEFTDGDGTQWRWESKHKTIDGKFLSAGFYGDSEKDKGLQTSQDARFYAM 72  
 QY 61 SASPEPSNKGQTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDIC 120  
 Db 73 SSRFDSFSNKGQTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDIC 132  
 QY 121 GPCTKKVHVIFNYKGNVLINKIRCKDDEFTHLYTLIVRPNTVEVKIDNSQVSGSLE 180  
 Db 133 GPCTKKVHVIFNYKGNVLINKIRCKDDEFTHLYTLIVRPNTVEVKIDNSQVSGSLE 192  
 RESULT 7  
 JH0795  
 calreticulin precursor - California sea hare  
 N:Alternate names: protein 407  
 C:Species: Aplysia californica (California sea hare)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: JH0795; B31409; F60977  
 R:Kennedy, T.E.; Kuhl, D.; Barzilai, A.; Sweatt, J.D.; Kandel, E.R.  
 Neuron 9, 1013-1024, 1992  
 A:Title: Long-term sensitization training in aplysia leads to an increase in calreticulin  
 A:Reference number: JH0795; MUID:93098937; PMID:1463604  
 A:Accession: JH0795  
 A:Molecule type: mRNA  
 A:Residues: 1-405 <KEN>  
 A:Cross-references: GB:S51239; NID:Q262053; PIDN:AA24569.1; PID:G262054  
 A:Experimental source: abdominal ganglion and atrial nervous system  
 R:Kennedy, T.E.; Gawinowicz, M.A.; Barzilai, A.; Kandel, E.R.; Sweatt, J.D.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 7008-7012, 1992  
 A:Title: Sequencing of proteins from two-dimensional gels by using in situ digestion and  
 A:Reference number: A94207; MUID:88320566; PMID:3413132  
 A:Accession: B31409  
 A:Molecule type: protein  
 A:Residues: 'X', 17-28, 'X', 30-31 <KE2>  
 R:Sweatt, J.D.; Kennedy, T.E.; Wager-Smith, K.; Gawinowicz, M.A.; Barzilai, A.; Karl, K.

Electrophoresis 10, 152-157, 1989  
 A:Title: Development of a database of amino acid sequences for proteins identified and i  
 A:Reference number: A60977; MUID:89276264; PMID:2731514  
 A:Accession: F60977  
 A:Molecule type: protein  
 A:Residues: 'X', 17-28, 'X', 30-31 <SWE>  
 C:Superfamily: calreticulin  
 C:Keywords: calcium binding; endoplasmic reticulum  
 F:1-15/Domain: signal sequence #status predicted <SIG>  
 F:16-405/Product: calreticulin #status experimental <MAT>  
 F:402-405/Region: endoplasmic reticulum retention signal

Query Match 82.8%; Score 804; DB 1; Length 405;  
 Best Local Similarity 79.4%; Pred. No. 7.7e-66;  
 Matches 143; Conservative 18; Mismatches 17; Indels 2; Gaps 1;  
 QY 1 EPAVYFKEQFLDGDGWTSMWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60  
 Db 16 DPTVYFKEEF--GDDWAERWVESKHKSGLCKFLVLTAGKFGYDAEKDKGLQTSQDARFYGL 73  
 QY 61 SASPEPSNKGQTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDIC 120  
 Db 74 SAKFPEPSNKGQTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDIC 133  
 QY 121 GPCTKKVHVIFNYKGNVLINKIRCKDDEFTHLYTLIVRPNTVEVKIDNSQVSGSLE 180  
 Db 134 GPCTKKVHVIFNYKGNVLINKIRCKDDEFTHLYTLIVRPNTVEVKIDNEKASGSDLE 193  
 RESULT 8  
 S29130  
 calreticulin (clone 8) - African clawed frog (fragment)  
 C:Species: Xenopus laevis (African clawed frog)  
 C:Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: S29130; T01068  
 R:Trevies, S.; Zorzato, F.; Pozzan, T.  
 Biochem. J. 287, 579-581, 1992  
 A:Title: Identification of calreticulin isoforms in the central nervous system.  
 A:Reference number: S29129; MUID:93074997; PMID:1445218  
 A:Accession: S29130  
 A:Molecule type: mRNA  
 A:Residues: 1-384 <TR>  
 A:Cross-references: UNIPROT:Q91711; EMBL:X67598  
 A:Accession: T01068  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Residues: 1-339, 'X'GR' <TR>  
 A:Cross-references: EMBL:X67598; NID:G64610; PIDN:CAA47867.1; PID:G64611  
 A:Experimental source: CNS  
 C:Superfamily: calreticulin  
 C:Keywords: glycoprotein  
 F:381-384/Region: endoplasmic reticulum retention signal  
 F:316/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 82.1%; Score 797; DB 2; Length 384;  
 Best Local Similarity 85.2%; Pred. No. 3.1e-65;  
 Matches 144; Conservative 12; Mismatches 13; Indels 0; Gaps 0;  
 QY 12 DGDGWTSMWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYALSASFEPFSNKG 71  
 Db 1 DGDGWTSMWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAMSRFSFSNKG 60  
 QY 72 QTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDICGGTCKKHVIF 131  
 Db 61 QTLVVOFTVKHEQNIDCGGYVKLPSPNSLDQTDHMGDSSEYINMGPDICGGTCKKHVIF 120  
 QY 132 NYKGNVLINKIRCKDDEFTHLYTLIVRPNTVEVKIDNSQVSGSLE 180  
 Db 121 QYKKNLQINKIRCKDDEFTHLYTLIVRPNTVEVKIDNSQVSGSLE 169  
 RESULT 9  
 S71343



calreticulin precursor - Korean frog  
C:Species: Rana rugosa (Korean frog)  
C>Date: 29-Jan-1998 #sequence\_revision 13-Feb-1998 #text\_change 20-Jun-2000  
C:Accession: S71343  
R:Yamamoto, S.; Nakamura, M.  
FEBS Lett. 387, 27-32, 1996  
A:title: Calnexin: its molecular cloning and expression in the liver of the frog, Rana x

A:Reference number: S71342; PMID:96234004; PMID:8654561  
A:Accession: S71343  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 1-419 <AM>  
A:Cross-references: EMBL:D78589; NID:g1514956; PIDN:BAA11425.1; PID:g1514957  
C:Superfamily: calreticuln  
C:Keywords: calcium binding; endoplasmic reticulum  
F:1-18/Domains: signal sequence #status predicted <SIG>  
F:19-419/Product: calreticulin #status predicted <SIG>  
F:205-213/Region: nuclear location signal  
F:415-418/Region: endoplasmic reticulum retention signal

Query Match            82.0%; Score 796; DB 2: Length 419;  
Best Local Similarity   80.6%; Pred. No. 4.3e-65;  
Matches 145; Conservative 19; Mismatches 16; Indels 0; Gaps 0;

QY         1 EPAYVKEPFLDGDGWTSWIESKHSDTGCKVLSSGKYGVDEEKDKGLQTODARFYAL 60  
            :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         19 KEAVYQEFGDGGANKERWIOQNHKSDDYGQFKLSAGKYGVDEEKDGKLQTSODAKFYAH 78  
            ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

QY         61 SASFEPFNKGQLTVQVTFVHEQNIDCGGYVKLFPSNSLDQTDHGDSEYNIMFGPDIC 120  
            :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         79 SAGFPAPSKNDKPLVVQVFSGVEHQNTDCGGYYVKLPSPSTLEQTMHGGESEYNIMEFGDIC 138  
            :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

QY         121 GPCTKKVHVIFNYKGNLVINKDIRCKDDSEFHLYTLIVRPONTYEYVKINDSNQVSZGSLE 180  
            :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:  
DB         139 GPCTKKVHVIFNYKGNLVINKDIRSKADVIHSHLYTLIVRPONTYEYVKINDSNKVESGNLE 198  
            :|::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|:

RESULT 10  
AS56637  
Calreticulin homolog precursor - fruit fly (*Drosophila melanogaster*)  
N:Alternate names: Ro/SS-A autoantigen/calreticulin homolog  
C:Species: Drosophila melanogaster  
C>Date: 11-Aug-1995 #sequence\_revision 11-Aug-1995 #text\_change 09-Jul-2004  
C:Accession: AS56637; A37158  
R:Smith, M.J.  
DNA Seq. 3, 247-250, 1992  
A>Title: Nucleotide sequence of a *Drosophila melanogaster* gene encoding a calreticulin H  
A:Reference number: AS56637; MUID:93208374; PMID:1296619  
A:Accession: AS56637  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-406 <SMI>  
A:Cross-references: UNIPROT:P29413; GB:X64461; NID:G7685; PIDN:CAA45791.1; PID:g7686  
A>Note: sequence extracted from NCBI backbone (NCBIN:I182874, NCBIPI:128275)  
R:MCCauliffe, D.P.; Zappi, E.; Lieber, T.S.; Michalak, M.; Sontheimer, R.D.; Capra, J.D.  
J. Clin. Invest. 86, 332-335, 1990  
A>Title: A human Ro/SS-A autoantigen is the homologue of calreticulin and is highly homo  
A:Reference number: A37158; MUID:90307991; PMID:2365822  
A:Accession: A37158  
A>Status: preliminary; nucleic acid sequence not shown; not compared with conceptual tra  
A:Molecule type: DNA  
A:Residues: 91-105, 'A', 107, 109-124, 182-183, 'L', 185-220 <MCC>  
C:Genetics:  
A:Gene: FlyBase:Crc  
A:Cross-references: FlyBase:Fgn0005585  
A:Introns: 65/1; 222/3  
C:Superfamily: calreticulin  
C:Keywords: calcium binding; endoplasmic reticulum  
F:1-17/Domains: signal sequence #status predicted <SIG>  
F:403-406/Region: endoplasmic reticulum retention signal

Query Match            72.0%; Score 699; DB 2: Length 406;  
Best Local Similarity   74.2%; Pred. No. 3.1e-56;

A:Title: Isolation and characterization of expression cDNA clones encoding antigens of calreticulin - barley (fragment)  
A:Reference number: A92769; MUID:88273584; PMID:2455736

A:Accession: A32507

A:Molecule type: mRNA

A:Residues: 1-336 <UNN>

A:Cross-references: UNIPROT:P11012

C:Superfamily: calreticulin

Query Match 55.9%; Score 542.5; DB 2; Length 336;  
Best Local Similarity 56.4%; Pred. No. 1.3e-41;  
Matches 98; Conservative 22; Mismatches 22; Indels 1; Gaps 1;

QY 39 FYGDEKDKGLQTSQDAFYALSASPE-PFSNKGQTLVQFTVHEQNDICGGYVKKLFP 97

Db 1 FYGDAVKDKGLTKTQDAKFSYIGAKFDKFSNKGSLVTFQSVKHEQDIDCGGYVKKLMA 60

QY 98 NSLDQTMHGDSEYNIMFGPDICGPGTKKVVHIFNYKGNVLINKDIRCKDDEFTHLYTL 157

Db 61 SDVNLEDHGETPHIMFGPDICGPGTKKVVHIFNYKGNVLINKDIRCKDDVFTHLYTL 120

QY 158 IVRPDNTYEVKIDNSOVESGSL 180

Db 121 IVNSDNTYEVQIDGERAESGELE 143

RESULT 13

T05703

calreticulin - barley (fragment)

C:Species: Hordeum vulgare (barley)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T05703

R:Chen, F.; Hayes, P.M.; Mulroony, D.; Pan, A.

Plant Cell 6, 835-843, 1994

A:Title: Identification and characterization of cDNA clones encoding plant calreticulin

A:Reference number: Z15422; MUID:94339696; PMID:7914763

A:Accession: T05703

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-412 <CHE>

A:Cross-references: UNIPROT:Q40040; EMBL:L27348; NID:G439585; PIDN:AAA32948.1; PID:G4395

A:Experimental source: cv. Morex, ovary

C:Genetics:

A:Gene: CRH1

A:Map position: 1M

C:Superfamily: calreticulin

C:Keywords: calcium binding

Query Match 55.5%; Score 539; DB 2; Length 412;

Best Local Similarity 56.4%; Pred. No. 1.3e-41;

Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHKSD---FGKVLSSGKFGYDEKDKGLQTSQDAFYAL 60

Db 20 VYFQEKF--EDGWSRWKSEWKKDENMAGWNHTSGKWHGDAE-DKGIQTSQDAFYAL 76

QY 61 SASPEFNSKQTLVQFTVHEQNDICGGYVKKLFPNSLDQTMHGDSEYNIMFGPDIC 120

Db 77 SAYPEFNSKDKTLVLQFTVHEQKLDCCGGYVKKLGGDVKQKFGGDTPTYGIMFGPDIC 136

QY 121 GPGTKKVVHIFNYKGNVLINKDIRCKDDEFTHLYTLVIRPDNTYEVKIDNSOVESGSL 179

Db 137 GYSTKKVHTLTNKGKNNHLIKDVPCTDQLSHVYTLIRPDATYSILIDNEEKQTGSI 195

RESULT 14

T05705

calreticulin - barley (fragment)

C:Species: Hordeum vulgare (barley)

C:Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004

C:Accession: T05705

R:Chen, F.; Hayes, P.M.; Mulroony, D.; Pan, A.

Plant Cell 6, 835-843, 1994

A:Title: Identification and characterization of cDNA clones encoding plant calreticulin

A:Reference number: Z15422; MUID:94339696; PMID:7914763

A:Accession: T05705

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-415 <CHE>

A:Cross-references: UNIPROT:Q40041; EMBL:L27349; NID:G439587; PIDN:AAA32949.1; PID:G4395

A:Experimental source: cv. Morex, ovary

C:Genetics:

A:Gene: CRH2

A:Map position: 2M

C:Superfamily: calreticulin

C:Keywords: calcium binding

Query Match 55.5%; Score 539; DB 2; Length 415;

Best Local Similarity 56.4%; Pred. No. 1.3e-41;

Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHKSD---FGKVLSSGKFGYDEKDKGLQTSQDAFYAL 60

Db 23 VYFQEKF--EDGWSRWKSEWKKDENMAGWNHTSGKWHGDAE-DKGIQTSQDAFYAL 79

QY 61 SASPEFNSKQTLVQFTVHEQNDICGGYVKKLFPNSLDQTMHGDSEYNIMFGPDIC 120

Db 80 SAYPEFNSKDKTLVLQFTVHEQKLDCCGGYVKKLGGDVKQKFGGDTPTYGIMFGPDIC 139

QY 121 GPGTKKVVHIFNYKGNVLINKDIRCKDDEFTHLYTLVIRPDNTYEVKIDNSOVESGSL 179

Db 140 GYSTKKVHTLTNKGKNNHLIKDVPCTDQLSHVYTLIRPDATYSILIDNEEKQTGSI 198

RESULT 15

S58170

calreticulin precursor - maize

C:Species: Zea mays (maize)

C:Date: 13-Jan-1996 #sequence\_revision 01-Mar-1996 #text\_change 13-Aug-1999

C:Accession: S58170; S49818

R:Dresselhaus, T.; Hagel, C.; Loerz, H.; Kranz, E.

submitted to the EMBL Data Library, July 1995

A:Description: Isolation of a cDNA encoding Calreticulin from in vitro zygotes of maize.

A:Reference number: S58170

A:Accession: S58170

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-421 <DRE>

A:Cross-references: EMBL:X89813; NID:G927571; PIDN:CAA61939.1; PID:G927572

R:Napier, R.M.; Trueman, S.; Henderson, J.; Boyce, J.M.; Hawes, C.R.; Ficker, M.D.; Ven

submitted to the EMBL Data Library, November 1994

A:Description: Purification and sequencing of calreticulin from maize and evidence for it

A:Reference number: S49818

A:Accession: S49818

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-421 <NAP>

A:Cross-references: EMBL:Z46772; NID:G577611; PIDN:CAA86728.1; PID:G577612

C:Genetics:

A:Gene: CRH1

C:Superfamily: calreticulin

C:Keywords: calcium binding

F:1-25/Domain: signal sequence #status predicted <SIG>

F:418-421/Region: endoplasmic reticulum retention signal

Query Match 55.2%; Score 536; DB 2; Length 421;

Best Local Similarity 56.4%; Pred. No. 2.6e-41;

Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHKSD---FGKVLSSGKFGYDEKDKGLQTSQDAFYAL 60

Db 27 VYFQEKF--EDGWSRWKSEWKKDENMAGWNHTSGKWHGDAE-DKGIQTSQDAFYAL 83

QY 61 SASPEFNSKQTLVQFTVHEQNDICGGYVKKLFPNSLDQTMHGDSEYNIMFGPDIC 120

Db 84 SAYPEFNSKDKTLVLQFVSVEHQKLDCCGGYVKKLGGDVKQKFGGDTPTYGIMFGPDIC 143

Qy 121 GPGTKKVHVFNYKGNVLIINKDIRCKDDEFTHLYTLIVRPNTYEVKIDNSQVESGSL 179  
Db 144 GYSTKKVHTILTKDGNHLIKDVFCTDQLTHVYTLIIRPDATYSILIDNEEKQTGSI 202

Search completed: October 15, 2004, 13:12:53  
Job time : 15 secs

**This Page Blank (uspto)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2004, 13:11:18 ; Search time 155 Seconds

(without alignments)  
416.589 Million cell updates/sec

Title: US-09-807-148-4

Perfect score: 971

Sequence: 1 EPAYVFKQFLDGDWTSRW.....PDNTYEVKIDNSQVSGSLE 180

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23sep04:\*

- 1: Geneseq1980s:\*
- 2: Geneseq1990s:\*
- 3: Geneseq2000s:\*
- 4: Geneseq2001s:\*
- 5: Geneseq2002s:\*
- 6: Geneseq2003as:\*
- 7: Geneseq2003bs:\*
- 8: Geneseq2004s:\*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	971	100.0	180	3	Aay92351 Human vas
2	971	100.0	387	7	Adg75706 Human pro
3	971	100.0	400	3	Aay92350 Recombina
4	971	100.0	417	1	Aap92276 60 KD Ro
5	971	100.0	417	3	Aay92349 Human MBP
6	971	100.0	417	5	Aau77712 Human cal
7	971	100.0	417	5	Aae24591 Human cal
8	971	100.0	417	5	Aae18851 Human cal
9	971	100.0	417	5	Abe82384 Human cal
10	971	100.0	417	6	Abj19766 Human MP2
11	971	100.0	417	6	Aag79824 Calreticu
12	971	100.0	417	6	Ada26337 Human cal
13	971	100.0	417	7	Add22407 HLA-B46 T
14	971	100.0	417	7	Ades6308 Human Pro
15	971	100.0	417	7	Ades6312 Human Pro
16	971	100.0	417	7	Adis7349 Human pvu
17	971	100.0	417	8	Adn03688 Antipepti
18	971	100.0	417	8	Ado19114 Human PRO
19	966	99.5	417	2	Aay00927 Calreticu
20	933	96.1	416	7	Ades6306 Rat Prote
21	933	96.1	416	7	Adf30443 Rat Prote
22	933	96.1	416	7	Adf30439 Rat angio
23	933	96.1	416	7	Adf30439 Rat angio
24	933	96.1	416	7	Adf30441 Rat angio
25	921	94.9	401	2	Aaw11156 Calreticu

26	700	72.1	403	2	Aaw04171 Flea calr
27	699	72.0	406	4	Abb64414 Drosophi
28	631	65.0	384	7	Adi57351 Human cal
29	631	65.0	385	3	Aab32385 Human sec
30	631	65.0	390	6	Ab071134 Novel hum
31	616.5	63.5	177	5	Abp42414 Human ova
32	541.5	55.8	336	2	Aar12312 Partial s
33	540	55.6	419	7	Abm74155 DNA clone
34	527	54.3	415	4	Aab66343 Castor be
35	527	54.3	415	4	Aab66341 Castor be
36	510	52.5	421	3	Agd24608 Arabidops
37	510	52.5	421	3	Agd47932 Arabidops
38	510	52.5	424	3	Agd4607 Arabidops
39	510	52.5	424	3	Agd7931 Arabidops
40	510	52.5	441	3	Agd30997 Arabidops
41	510	52.5	444	3	Agd30996 Arabidops
42	494	50.9	420	5	Abb04656 Maize cal
43	486	50.1	428	7	Abm74288 DNA clone
44	327	33.7	61	3	Aay92352 Recombina
45	318	32.7	60	3	Aay92354 Recombina

ALIGNMENTS

RESULT 1

AAY92351

ID AAY92351 standard; protein; 180 AA.

XX AC AAY92351;

XX DT 10-AUG-2000 (first entry)

XX DE Human vasostatin (calreticulin N-terminal 180 amino acids).

XX KW MBP-calreticulin; maltose binding protein; vasostatin; N-terminal; angiogenesis; inhibition; endothelial cell; anti-angiogenic; neuroprotective; antidiabetic; cytostatic; dermatological; hepatic; immunosuppressive; antiinflammatory; anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.

XX OS Homo sapiens.

XX OS Synthetic.

XX FN WO200020577-A1.

XX PD 13-APR-2000.

XX PF 05-OCT-1999; 99WO-US023240.

XX PR 06-OCT-1998; 98US-0103438P.

XX PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

XX PI Tosato G, Pike SE, Yao L;

XX WPI; 2000-303767/26.

XX PT Inhibiting endothelial cell growth and angiogenesis using calreticulin, useful for suppressing tumor growth.

XX PS Claim 4; Page 82; 99pp; English.

XX CC A novel method of inhibiting endothelial cell growth comprises contacting the cells with calreticulin (or its fragments/variants). Fragments of calreticulin causes at least 40% inhibition of angiogenesis, tumor growth and/or endothelial cell growth (claimed). The method may be used for inhibiting angiogenesis in a patient. The angiogenesis is associated with a disease other than a tumor that is associated with neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiodiomas, immune inflammation, atherosclerosis, excessive wound repair, retinal neovascularization, macular degeneration, corneal graft rejection, contact lens overwear,

CC Crohn's disease, non-immune inflammation, rheumatoid arthritis, systemic  
 CC lupus erythematosus, thyroiditis, Goodpasture's syndrome, systemic  
 CC vasculitis, scleroderma, Sjorgen's syndrome, sarcoidosis and primary  
 CC biliary cirrhosis). The method may also be used for treating/inhibiting  
 CC tumor growth especially Kaposi's sarcoma (claimed)  
 XX Sequence 180 AA;  
 SQ

Query Match 100.0%; Score 971; DB 3; Length 180;  
 Best Local Similarity 100.0%; Pred No. 4.4e-99;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVLSGKFGYDDEKDKGLQTSQDARFYAL 60  
 Db 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVLSGKFGYDDEKDKGLQTSQDARFYAL 60

QY 61 SASFEFPSNKGOTLVVQFTVKHEQNIIDCGGVVVKLFPSLSDOTDMHGDSEYNIMFGPDIC 120  
 Db 61 SASFEFPSNKGOTLVVQFTVKHEQNIIDCGGVVVKLFPSLSDOTDMHGDSEYNIMFGPDIC 120

QY 121 GPGTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLTVRPDNTYEVKIDNSQVESGSLE 180  
 Db 121 GPGTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLTVRPDNTYEVKIDNSQVESGSLE 180

RESULT 2  
 ADG75706  
 ID ADG75706 standard; protein; 387 AA.  
 AC ADG75706;  
 XX  
 DT 11-MAR-2004 (first entry)  
 DE Human protein modification and maintenance molecule polypeptide SeqID30.  
 XX  
 KW human; protein modification and maintenance molecule; PMMW; kinase;  
 KW phosphatase; protease; protease inhibitor; isomerase; transferase;  
 KW molecular chaperone; gastrointestinal; cardiovascular; autoimmune;  
 KW inflammatory; cell proliferative; developmental; epithelial;  
 KW neurological; reproductive disorder; enzyme;  
 KW single nucleotide polymorphism; SNP.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003083084-A2.  
 XX  
 PD 09-OCT-2003.  
 XX  
 PF 26-MAR-2003; 2003WO-US009774.  
 XX  
 PR 29-MAR-2002; 2002US-0368686P.  
 PR 03-MAY-2002; 2002US-0377489P.  
 PR 03-MAY-2002; 2002US-0378205P.  
 XX  
 PA (INCY-) INCYTE CORP.  
 XX  
 PI Lal PG, Tang YT, Lee SY, Griffin JA, Elliott VS, Khare R;  
 PI Ramkumar J, Richardson TW, Ison CH, Hafalia AJA, Kable AE;  
 PI Swarnakar A, Chang H, Bhatia UG, Burrill JD, Lee S, Blake JJ, Ho A;  
 PI Zheng W, Chawla NK, Marquis JP, Tran UK, Emerling BV, Mason PM;  
 PI Becha SD, Warren BA, Au-Young JK, Lee EA, Gietzen KU, Jiang X;  
 PI Jackson AA, Baughn MR, Wilson AD, Jin P, Bulloch SA;  
 XX  
 DR WPI; 2003-812541/76.  
 DR N-PSDB; ADG75776.  
 XX  
 PT New human protein modification and maintenance molecules (PMMW)  
 PT polypeptide, useful for preparing a composition for treating a disease  
 PT associated with decreased expression or overexpression of PMMW e.g.,  
 cancer.  
 XX  
 PS Claim 1; SEQ ID NO 30; 500pp; English.  
 PS  
 XX

CC This invention relates to novel isolated polynucleotides that encode  
 CC human protein modification and maintenance molecules (PMMW).  
 CC Specifically, it refers to key enzymes and proteins such as kinases,  
 CC phosphatases, proteases, protease inhibitors, isomerases, transferases  
 CC and molecular chaperones that modify and maintain protein molecules by co-  
 CC ordinating their function, conformation, stabilisation and degradation.  
 CC The present invention describes antibodies, agonists and antagonists (as  
 CC well as single nucleotide polymorphisms) that can be useful for the  
 CC diagnosis, treatment and prevention of disorders associated with the  
 CC aberrant expression of PMMW, which includes gastrointestinal,  
 CC cardiovascular, autoimmune/ inflammatory, cell proliferative,  
 CC developmental, epithelial, neurological and reproductive disorders.  
 CC Furthermore, the invention relates to the effect of exogenous compounds  
 CC on the expression of PMMWs. This polypeptide is a human PMMW protein of  
 CC the invention.  
 XX Sequence 387 AA;  
 SQ

Query Match 100.0%; Score 971; DB 7; Length 387;  
 Best Local Similarity 100.0%; Pred No. 1.3e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVLSGKFGYDDEKDKGLQTSQDARFYAL 60  
 Db 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVLSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASFEFPSNKGOTLVVQFTVKHEQNIIDCGGVVVKLFPSLSDOTDMHGDSEYNIMFGPDIC 120  
 Db 78 SASFEFPSNKGOTLVVQFTVKHEQNIIDCGGVVVKLFPSLSDOTDMHGDSEYNIMFGPDIC 137

QY 121 GPGTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLTVRPDNTYEVKIDNSQVESGSLE 180  
 Db 138 GPGTKKVHVIFNYKGNVLINKDIRCKDDEFTHLTYLTVRPDNTYEVKIDNSQVESGSLE 197

RESULT 3  
 AAY92350  
 ID AAY92350 standard; protein; 400 AA.  
 XX  
 AC AAY92350;  
 XX  
 DT 10-AUG-2000 (first entry)  
 XX  
 DE Recombinant human MBP-calreticulin.  
 XX  
 KW MBP-calreticulin; maltose binding protein; angiogenesis; inhibition;  
 KW endothelial cell; anti-angiogenic; neuroprotective; antidiabetic;  
 KW cytostatic; dermalogical; immunosuppressive; antiinflammatory; hepatic;  
 KW anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200020577-A1.  
 XX  
 PD 13-APR-2000.  
 XX  
 PF 05-OCT-1999; 99WO-US023240.  
 XX  
 PR 06-OCT-1998; 98US-0103438P.  
 XX  
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.  
 XX  
 PI Tosato G, Pike SE, Yao L;  
 XX  
 DR WPI; 2000-303767/26.  
 DR N-PSDB; AAA09346, AAA09347.  
 XX  
 PT Inhibiting endothelial cell growth and angiogenesis using calreticulin,  
 PT useful for suppressing tumor growth.  
 XX  
 PS Claim 4; Page 80-81; 99pp; English.  
 PS  
 XX Recombinant human MBP-calreticulin comprises the sequence of human

calreticulin (see AAY92349) without the 17 N-terminal amino acids. A novel method of inhibiting endothelial cell growth comprises contacting the cells with calreticulin (or its fragments/variants). Fragments of calreticulin causes at least 40% inhibition of angiogenesis, tumor growth and/or endothelial cell growth (claimed). The method may be used for inhibiting angiogenesis in a patient. The angiogenesis is associated with a disease other than a tumor that is associated with neovascularization (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular glaucoma, psoriasis, angiodiomas, immune inflammation, atherosclerosis, excessive wound repair, retinal neovascularization, macular degeneration, corneal graft rejection, contact lens overwear, Crohn's disease, non-immune inflammation, rheumatoid arthritis, systemic lupus erythematosus, thyroiditis, Goodpasture's syndrome, systemic vasculitis, scleroderma, Sjorgen's syndrome, sarcoidosis and primary biliary cirrhosis). The method may also be used for treating/inhibiting tumor growth especially Kaposi's sarcoma (claimed)

Query Match 100.0%; Score 971; DB 3; Length 400;  
Best Local Similarity 100.0%; Pred. No. 1.3e-98;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVFKQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDSEKDKGLQTSQDARFYAL 60  
DB 1 EPAYVFKQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDSEKDKGLQTSQDARFYAL 60  
QY 61 SASFEFNSKGTQVLVQFTVKHEQNIIDCGGVYKLPNSLDQTMHGDSEYNIMFGPDIC 120  
DB 61 SASFEFNSKGTQVLVQFTVKHEQNIIDCGGVYKLPNSLDQTMHGDSEYNIMFGPDIC 120  
QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 180  
DB 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 180

RESULT 4  
AAP92276  
ID AAP92276 standard; protein; 417 AA.  
AC AAP92276;  
XX  
XX 25-MAR-2003 (revised)  
DT 23-FEB-1990 (first entry)  
XX  
XX 60 kD Ro (Rc/SSA) antigen.  
XX  
XX Sjorens syndrome; systemic lupus erythematosus.  
XX  
XX Synthetic.  
XX  
XX WO8909273-A.  
XX  
XX 05-OCT-1989.  
XX  
XX 22-MAR-1989; 89WO-US001213.  
XX  
XX 22-MAR-1988; 88UG-00171634.  
XX  
XX (TEXA ) UNIV TEXAS SYSTEM.  
XX  
XX Sontheimer RD, Lieu TS, Capra JD, Mccaulliffe DP;  
XX  
XX WPI; 1989-309537/42.  
XX  
XX N-PSDB; AAP92276.  
XX  
XX DNA sequences encoding antigenic epitope(s) of Ro 60 KD auto-antigen - used in immunoassays to detect rheumatic disease.  
XX  
XX Disclosure; Fig 2; 88pp; English.  
XX  
XX Synthetic peptides corresp. to an epitopic core of Ro antigen are expressed recombinantly to detect autoantibodies, for identification of

autoimmune diseases. These epitopes are AAs 24-36, 23-36, 188-209, or 241-255. The peptides may be substd. for ribonucleoprotein particle antigens. (Updated on 25-MAR-2003 to correct PA field.) (Updated on 25-MAR-2003 to correct PI field.)

Query Match 100.0%; Score 971; DB 1; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVFKQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDSEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVFKQFLDGDGWTSEWIESKHKSDFGKFLVSSGKFGYGDSEKDKGLQTSQDARFYAL 77  
QY 61 SASFEFNSKGTQVLVQFTVKHEQNIIDCGGVYKLPNSLDQTMHGDSEYNIMFGPDIC 120  
DB 78 SASFEFNSKGTQVLVQFTVKHEQNIIDCGGVYKLPNSLDQTMHGDSEYNIMFGPDIC 137  
QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 180  
DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYLIVRPDNTYEVKIDNSQVESGSLE 197

RESULT 5  
AAY92349  
ID AAY92349 standard; protein; 417 AA.  
AC AAY92349;  
XX  
XX 10-AUG-2000 (first entry)  
DT  
XX Human MBP-calreticulin.  
XX  
XX MBP-calreticulin; maltose binding protein; angiogenesis; inhibition; endothelial cell; anti-angiogenic; neuroprotective; antidiabetic; cytostatic; dermatologic; immunosuppressive; antiinflammatory; hepatic; anti-atherosclerotic; gastrointestinal; anti-arthritis; ophthalmic.  
XX  
XX Homo sapiens.  
XX  
XX Key  
FH Peptide 1. 17  
FT /label= signal\_peptide  
FT Protein 18  
FT /label= mature\_protein  
XX  
XX WO200020577-A1.  
XX  
XX 13-APR-2000.  
XX  
XX 05-OCT-1999; 99WO-US023240.  
XX  
XX 06-OCT-1998; 98US-0103438P.  
XX  
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
XX  
XX Tosato G, Pike SE, Yao L;  
XX  
XX WPI; 2000-303767/26.  
XX  
XX N-PSDB; AAA09346, AAA09347.  
XX  
XX Inhibiting endothelial cell growth and angiogenesis using calreticulin, useful for suppressing tumor growth.  
XX  
XX Disclosure; Page 79-80; 99pp; English.  
XX  
XX A novel method of inhibiting endothelial cell growth comprises contacting the cells with calreticulin (or its fragments/variants). Fragments of calreticulin causes at least 40% inhibition of angiogenesis, tumor growth and/or endothelial cell growth (claimed). The method may be used for inhibiting angiogenesis in a patient. The angiogenesis is associated with a disease other than a tumor that is associated with neovascularization

CC (e.g. diabetic neuropathy, retrolental fibroplasia, trachoma, neovascular  
 CC glaucoma, psoriasis, angiofibromas, immune inflammation.  
 CC atherosclerosis, excessive wound repair, retinal neovascularization,  
 CC macular degeneration, corneal graft rejection, contact lens overwear,  
 CC Crohn's disease, non-immune inflammation, rheumatoid arthritis, systemic  
 CC lupus erythematosus, thyroiditis, Goodpasture's syndrome, systemic  
 CC vasculitis, scleroderma, Sjorgen's syndrome, sarcoidosis and primary  
 CC biliary cirrhosis). The method may also be used for treating/inhibiting  
 CC tumor growth especially Kaposi's sarcoma (claimed)  
 XX  
 SQ Sequence 417 AA;

Query Match 100.0%; Score 971; DB 3; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60  
 Db 18 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 77  
 QY 61 SASFEFPSNKGQTLVVQFTVKHEQNDICGGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 120  
 Db 78 SASFEFPSNKGQTLVVQFTVKHEQNDICGGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 137  
 QY 121 GPGTKKVHVIIFYNKGKQNLINKDIRCKDDEFTHLTYLIVRPDNTVEVKIDNSQVESGSLE 180  
 Db 138 GPGTKKVHVIIFYNKGKQNLINKDIRCKDDEFTHLTYLIVRPDNTVEVKIDNSQVESGSLE 197

RESULT 6  
 AAU7712  
 ID AAU7712 standard; protein; 417 AA.  
 XX  
 AC AAU7712;  
 XX  
 DT 05-JUN-2002 (first entry)  
 DE Human calreticulin (CRT).  
 KW Calreticulin; CRT; endoplasmic reticulum chaperone polypeptide;  
 KW cytotaxtic; vaccine; human papillomavirus 16; HPV 16; E7; DNA vaccine;  
 KW enhanced antigen-specific immune response; cytotoxic T lymphocyte;  
 KW tumour; cancer; cervical cancer.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200212281-A2.  
 XX  
 PD 14-FEB-2002.  
 XX  
 PF 02-AUG-2001; 2001WO-US024134.  
 XX  
 PR 03-AUG-2000; 2000US-0222902P.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Wu T, Hung C;  
 XX  
 DR WPI; 2002-257463/30.  
 DR N-PSDB; ABK11662.  
 XX

New nucleic acids encoding a fusion polypeptide comprising an endoplasmic  
 reticulum chaperone polypeptide linked to an antigenic polypeptide,  
 useful as a vaccine for inducing antigen-specific immune responses.

Disclosure; Page 27; 71pp; English.

The invention describes a nucleic acid molecule (1) encoding a fusion  
 polypeptide comprising a first polypeptide domain comprising an  
 endoplasmic reticulum chaperone polypeptide e.g. calreticulin (CRT) and a  
 second polypeptide domain comprising at least one antigenic peptide e.g.  
 Human papillomavirus 16 (HPV 16) E7. The nucleic acid is useful as a  
 vaccine (DNA vaccine) for inducing enhanced antigen-specific immune

CC responses, particularly those mediated by cytotoxic T lymphocytes. The  
 CC nucleic acid and compositions comprising the nucleic acid is also useful  
 CC for inhibiting the growth of tumours and cancers e.g. cervical cancer.  
 CC This is the amino acid sequence of the human calreticulin (CRT), an  
 CC endoplasmic reticulum protein used in the creation of a DNA vaccine  
 XX  
 SQ Sequence 417 AA;

Query Match 100.0%; Score 971; DB 5; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60  
 Db 18 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 77  
 QY 61 SASFEFPSNKGQTLVVQFTVKHEQNDICGGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 120  
 Db 78 SASFEFPSNKGQTLVVQFTVKHEQNDICGGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 137  
 QY 121 GPGTKKVHVIIFYNKGKQNLINKDIRCKDDEFTHLTYLIVRPDNTVEVKIDNSQVESGSLE 180  
 Db 138 GPGTKKVHVIIFYNKGKQNLINKDIRCKDDEFTHLTYLIVRPDNTVEVKIDNSQVESGSLE 197

RESULT 7  
 AAE24591  
 ID AAE24591 standard; protein; 417 AA.  
 XX  
 AC AAE24591;  
 XX  
 DT 04-OCT-2002 (first entry)  
 DE Human calreticulin protein.  
 XX  
 KW Human; calreticulin; antisense compound; hyperproliferative disorder;  
 KW cancer; autoimmune disease; viral infection; cardiovascular disease;  
 KW antisense therapy; cytostatic; immunosuppressive; virucide.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200236743-A2.  
 XX  
 PD 10-MAY-2002.  
 XX  
 PF 30-OCT-2001; 2001WO-US049045.  
 XX  
 PR 30-OCT-2000; 2000US-00702327.  
 XX  
 PA (ISIS-) ISIS PHARM INC.  
 XX  
 PI Bennett CF, Cowser LM;  
 XX  
 DR WPI; 2002-479759/51.  
 DR N-PSDB; AAD39469.  
 XX

Novel antisense compound targeted to nucleic acid encoding calreticulin,  
 useful for treating a human having disease or condition associated with  
 calreticulin e.g. cancer, viral infection, autoimmune disease.

Disclosure; Page 88-90; 109pp; English.

The invention relates to antisense compounds, compositions and methods  
 for modulating the expression of calreticulin. The compositions comprise  
 antisense compounds, particularly antisense oligonucleotides, targeted  
 to nucleic acids encoding calreticulin. The antisense compound is useful  
 for inhibiting the expression of calreticulin in human cells or tissues.  
 It is also useful for treating a human having a disease or condition  
 associated with calreticulin, e.g., hyperproliferative disorder e.g.  
 cancer, autoimmune disease, viral infection or cardiovascular disease, by  
 inhibiting expression of calreticulin. It is useful for diagnostics,  
 therapeutics, prophylaxis and as research reagents and kits. It is also  
 used in antisense therapy. The present sequence is human calreticulin



CC protein. This sequence is used in the exemplification of the invention  
 XX  
 SQ Sequence 417 AA;  
 Query Match 100.0%; Score 971; DB 5; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60  
 DB 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 77  
 QY 61 SASPEPFSNKGQTLVQVFTVKHEQNIIDCGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 120  
 DB 78 SASPEPFSNKGQTLVQVFTVKHEQNIIDCGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 137  
 QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180  
 DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 197  
 RESULT 8  
 ID AAE18851 standard; protein; 417 AA.  
 XX  
 AC AAE18851;  
 DT 17-MAY-2002 (first entry)  
 XX  
 DE Human calreticulin protein.  
 XX  
 KW Human; prostate cancer; calreticulin; T1D-1 protein; TRAITS protein;  
 KW androgen action pathway; cell proliferation; kidney cancer; lymphoma;  
 KW epithelium-derived carcinoma; leukaemia; vaccine; gene therapy;  
 KW cytostatic; U19 protein.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Domain 98..170  
 FT Region /label= N-terminal\_domain  
 FT Region 98..103  
 FT Region /label= Alpha\_helix  
 FT Region 149..154  
 FT Domain /label= Alpha\_helix  
 FT Domain 171..285  
 FT Domain /note= "Proline-rich domain (P domain)"  
 FT Domain 286..397  
 FT Domain /label= C-terminal\_domain  
 XX  
 PN WC200206327-A2.  
 XX  
 PD 24-JAN-2002.  
 XX  
 PF 17-JUL-2001; 2001WO-USC22357.  
 XX  
 PR 17-JUL-2000; 2000US-0218761P.  
 PR 16-JUL-2001; 2001US-00906393.  
 XX  
 PA (NOUN ) UNIV NORTHWESTERN.  
 XX  
 PI Wang Z, Xiao W;  
 XX  
 DR WPI; 2002-179780/23.  
 DR N-PSDB; AAD29931.  
 XX  
 PT Identifying a subject that is likely to have aggressive form of prostate  
 PT cancer, involves comparing calreticulin levels in prostate specimen of  
 PT the subject and in benign prostatic epithelial cells of the same subject.  
 XX  
 PS Disclosure; Page 146-148; 148pp; English.  
 XX  
 PT The present invention relates to methods of distinguishing aggressive

CC forms of prostate cancer from non-aggressive forms. The method involves  
 CC comparing the level of calreticulin in prostate specimen and in benign  
 CC prostatic epithelial cells of a subject. The invention particularly  
 CC relates to two proteins, namely calreticulin and T1D-1 (TRAITS; U19) that  
 CC are down-regulated in aggressive forms of prostate cancer but not in  
 CC slowly progressing prostate cancer. They play important roles in the part  
 CC of androgen action pathway that suppresses cell proliferation and/or  
 CC prevents prostate cancer. The method is useful for identifying a subject  
 CC who is likely to have an aggressive form of prostate cancer. The  
 CC invention further relates to a method of identifying a subject with a  
 CC slow growing form of prostate cancer. T1D-1 sequences are useful for  
 CC treating cancers such as epithelium-derived carcinomas, kidney cancers,  
 CC lymphomas, leukaemias and prostate cancers. Sequences of the invention  
 CC are used as vaccines and in gene therapy. The present sequence is human  
 CC calreticulin protein  
 XX  
 SQ Sequence 417 AA;  
 Query Match 100.0%; Score 971; DB 5; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60  
 DB 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 77  
 QY 61 SASPEPFSNKGQTLVQVFTVKHEQNIIDCGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 120  
 DB 78 SASPEPFSNKGQTLVQVFTVKHEQNIIDCGGVYKLPFNSLDQTMHGDSEYNIMFGPDIC 137  
 QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180  
 DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 197  
 RESULT 9  
 ID ABB82384 standard; protein; 417 AA.  
 XX  
 AC ABB82384;  
 DT 08-JAN-2003 (first entry)  
 XX  
 DE Human calreticulin (CRT) protein.  
 XX  
 KW Immunogenicity; cytostatic; virucide; protozoacide; antibacterial; CTL;  
 KW hepatotropic; anti-HIV; vaccine. cytotoxic T lymphocyte; tumour; CRT;  
 KW calreticulin; human.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200274920-A2.  
 XX  
 PD 26-SEP-2002.  
 XX  
 PF 18-MAR-2002; 2002WO-US008033.  
 XX  
 PR 16-MAR-2001; 2001US-0276854P.  
 XX  
 PA (UYJO ) UNIV JOHNS HOPKINS.  
 XX  
 PI Wu T, Hung C;  
 XX  
 DR WPI; 2002-740856/80.  
 DR N-PSDB; ABV73173.  
 XX  
 PT New nucleic acid molecule encoding an antigenic fusion polypeptide useful  
 PT as vaccines for enhancing or inducing immune responses, primarily  
 PT cytotoxic T lymphocytes (CTL) responses to specific antigens such as  
 PT tumor or viral antigens.  
 XX  
 PS Disclosure; Page 27; 93pp; English.  
 XX

CC The invention relates to a nucleic acid molecule that encodes a fusion  
 CC polypeptide, comprising a first nucleic acid sequence encoding a  
 CC polypeptide that comprises at least one immunogenicity-potentiating  
 CC polypeptide, optionally fused in frame with the nucleic acid, a linker  
 CC nucleic acid encoding a linker peptide, and a nucleic acid that is linked  
 CC in frame to them, and that encodes an antigenic peptide or polypeptide.  
 CC The nucleic acid molecule, polypeptides and vectors are useful as  
 CC vaccines for enhancing immune responses, primarily cytotoxic T  
 CC lymphocytes (CTL) responses to specific antigens such as tumour or viral  
 CC antigens, and for inhibiting growth or preventing re-growth of a tumour.  
 CC The packaging cell line is useful for generating re-growth of a tumour.  
 CC Particles without contamination from replicon-competent virus. The  
 CC pathogenic organisms include viruses such as human papilloma virus (HPV),  
 CC hepatitis B virus, hepatitis C virus, human immunodeficiency virus,  
 CC Epstein Barr virus and herpes simplex virus, intracellular parasites such  
 CC as malaria, and bacteria that grow intracellularly such as mycobacteria  
 CC and listeria. The present sequence represents a human calreticulin (CRT)  
 XX  
 XX Sequence 417 AA;

Query Match 100.0%; Score 971; DB 5; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKPVLSGGKFGYDDEKDKLQTSQDARFYAL 60  
 Db 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKPVLSGGKFGYDDEKDKLQTSQDARFYAL 77

QY 61 SASPEFPFNKGGTLLVQFTVKEHQNIDCGGKGVVYKLPFNSLDQTDHMGDSEYNIMFGPDIC 120  
 Db 78 SASPEFPFNKGGTLLVQFTVKEHQNIDCGGKGVVYKLPFNSLDQTDHMGDSEYNIMFGPDIC 137

QY 121 GPGTKKXHVHVFNYKGNVLLINKDIRCKDDETHLYTLIVRPDNTVEVKIDNSQVESGSLE 180  
 Db 138 GPGTKKXHVHVFNYKGNVLLINKDIRCKDDETHLYTLIVRPDNTVEVKIDNSQVESGSLE 197

RESULT 10  
 ABU19766  
 ID ABU19766 standard; protein; 417 AA.  
 XX  
 AC ABU19766;  
 XX  
 DT 03-APR-2003 (first entry)  
 DE Human MP21 protein CRTC SEQ ID No 46.  
 XX  
 KW Cytostatic; p21 pathway modulating agent; cancer; angiogenic; apoptotic;  
 KW cell proliferation disorder; MP21.  
 XX  
 OS Homo sapiens.  
 XX  
 FN WO2003006990-A1.  
 XX  
 PD 23-JAN-2003.  
 XX  
 PF 10-JUL-2002; 2002WO-US021549.  
 XX  
 PR 12-JUL-2001; 2001US-0305017P.  
 PR 10-OCT-2001; 2001US-0328491P.  
 PR 15-FEB-2002; 2002US-0357452P.  
 XX  
 PA (EXEL-) EXELIXIS INC.  
 XX  
 PI Friedman L, Plowman GD, Belvin M, Li D, Funke RP;  
 XX  
 DR WPI; 2003-221779/21.  
 DR N-PSDB; ABT17040.  
 XX  
 PT Identifying candidate p21 pathway modulator, by contacting an assay  
 PT system having modifiers of p21 polypeptide or gene with a test agent to  
 PT provide a reference activity in system and detecting test agent-biased  
 PT activity.

XX Example; Page 153-155; 199pp; English.  
 XX  
 CC The invention relates to a novel method for identifying a candidate p21  
 CC pathway modulating agent. The novel method comprises contacting an assay  
 CC system, comprising a purified MP21 polypeptide (modifier of p21) or  
 CC nucleic acid, with a test agent under conditions, so that but for the  
 CC presence of a test agent, the assay system provides a reference activity  
 CC and detection of test agent-biased activity of the assay system. The  
 CC novel method of the invention is useful for identifying a candidate p21  
 CC pathway modulating agent. The invention also includes a method for  
 CC modulating the p21 pathway of a cell, and a method for diagnosing a  
 CC disease e.g. cancer in a patient. The identified modulators are useful in  
 CC diagnosis, therapy and pharmaceutical development. The modulators are  
 CC useful in a variety of diagnostic and therapeutic applications including  
 CC angiogenic, apoptotic and cell proliferation disorders. This sequence  
 CC represents an MP21 protein of the invention  
 XX  
 XX Sequence 417 AA;

Query Match 100.0%; Score 971; DB 6; Length 417;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKPVLSGGKFGYDDEKDKLQTSQDARFYAL 60  
 Db 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKPVLSGGKFGYDDEKDKLQTSQDARFYAL 77

QY 61 SASPEFPFNKGGTLLVQFTVKEHQNIDCGGKGVVYKLPFNSLDQTDHMGDSEYNIMFGPDIC 120  
 Db 78 SASPEFPFNKGGTLLVQFTVKEHQNIDCGGKGVVYKLPFNSLDQTDHMGDSEYNIMFGPDIC 137

QY 121 GPGTKKXHVHVFNYKGNVLLINKDIRCKDDETHLYTLIVRPDNTVEVKIDNSQVESGSLE 180  
 Db 138 GPGTKKXHVHVFNYKGNVLLINKDIRCKDDETHLYTLIVRPDNTVEVKIDNSQVESGSLE 197

RESULT 11  
 AAG79824  
 ID AAG79824 standard; protein; 417 AA.  
 XX  
 AC AAG79824;  
 XX  
 DT 16-APR-2003 (first entry)  
 DE Calreticulin.  
 XX  
 KW Motif; SE; signal transduction; Alzheimer's disease; neuronal cell;  
 KW calreticulin; arrhythmia; idiopathic nephritic syndrome;  
 KW non-autoimmune hyperthyroidism; obesity; polycystic kidney disease;  
 KW cancer; asthma; atopic dermatitis; psoriasis; rheumatoid arthritis;  
 KW systemic lupus erythematosus; infectious disease; atherosclerosis;  
 KW inflammatory bowel disease; osteoarthritis; septic shock;  
 KW ischaemia-reperfusion injury.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200299061-A2.  
 XX  
 PD 12-DEC-2002.  
 XX  
 PF 04-JUN-2002; 2002WO-US017536.  
 XX  
 PR 04-JUN-2001; 2001US-0295691P.  
 PR 03-JUN-2002; 2002US-00161959.  
 XX  
 PA (UNMI ) UNIV MICHIGAN.  
 XX  
 PI Holoshitz J, Ling S;  
 XX  
 DR WPI; 2003-156853/15.

PT Treating diseases associated with signal transduction aberrations, e.g.  
PT Alzheimer's disease, arrhythmia or rheumatoid arthritis, comprises  
PT administering a preparation comprising an SE- or an SE motif-containing  
PT peptide.  
XX  
PS Disclosure; Fig 14; 97pp; English.  
XX  
XX This sequence represents human calreticulin. This protein was used in the  
CC method of the invention for treating diseases associated with signal  
CC transduction aberrations. The method comprises: (a) providing a subject  
CC with one or more symptoms of Alzheimer's disease and, optionally, a  
CC plurality of neuronal cells expressing calreticulin, and a preparation  
CC comprising an shared epitope (SE) - containing peptide or a peptide which  
CC binds the calreticulin; and (b) administering the preparation to the  
CC subject. The inventive method is useful for treating diseases associated  
CC with signal transduction aberrations, such as Alzheimer's disease,  
CC arrhythmia, idiopathic nephritic syndrome, non-autoimmune  
CC hyperthyroidism, obesity, polycystic kidney disease, cancer, asthma,  
CC atopic dermatitis, psoriasis, rheumatoid arthritis, systemic lupus  
CC erythematosus, infectious diseases, inflammatory bowel disease,  
CC osteoarthritis, septic shock, atherosclerosis, congestive heart failure,  
CC insulin-resistance syndrome, and ischaemia-reperfusion injury  
XX  
SQ Sequence 417 AA;

Query Match 100.0%; Score 971; DB 6; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60  
Db 18 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 77  
QY 61 SASPEPSNKGOTLVQVFTVKHEQNTDCGGYVKLPFNSLDQTDHGDSEYNIMFGPDIC 120  
Db 78 SASPEPSNKGOTLVQVFTVKHEQNTDCGGYVKLPFNSLDQTDHGDSEYNIMFGPDIC 137  
QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 180  
Db 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 197

RESULT 12  
ADA26337  
ID ADA26337 standard; protein; 417 AA.  
AC ADA26337;  
XX  
XX 20-NOV-2003 (first entry)  
DT  
XX Human calreticulin protein SEQ ID NO:14.  
DE  
XX stress response protein; immune response; tumour;  
KW tumour growth inhibition; metastasis; cytostatic; vaccine; gene therapy;  
KW human; calreticulin.  
XX  
XX Homo sapiens.  
XX WO2003068941-A2.  
XX  
XX 21-AUG-2003.  
XX  
XX 13-FEB-2003; 2003WO-US004631.  
XX  
XX 13-FEB-2002; 2002US-0356293P.  
XX  
XX (UYDU-) UNIV DUKE.  
XX  
XX Nicchita C, Baker-Lepain J;  
XX  
XX WPI; 2003-679635/64.  
XX  
XX N-PSDB; ADA26336.  
XX

PT New Hsp60, Hsp70, Hsp90 or calreticulin polypeptide, useful for preparing  
PT a composition for inhibiting tumor growth or metastasis in a subject.  
XX  
PS Disclosure; Page 137-138; 150pp; English.  
XX  
XX The present invention describes a recombinant stress response polypeptide  
CC (I) free of an antigen-binding domain, where the recombinant stress  
CC response polypeptide comprises an extracellularly transported polypeptide  
CC eliciting an immune response in a subject; (2) eliciting an immune  
CC response in a subject; and (3) inhibiting tumour growth or metastasis in  
CC a subject. (I) has cytostatic activity, and can be used for preparing a  
CC in gene therapy. The polypeptide (I) can be used for preparing a  
CC composition for inhibiting tumour growth or metastasis in a subject. The  
CC present sequence represents human calreticulin, which is a stress  
CC response polypeptide given in the exemplification of the present  
CC invention.  
XX  
SQ Sequence 417 AA;

Query Match 100.0%; Score 971; DB 6; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 60  
Db 18 EPAYVFEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL 77  
QY 61 SASPEPSNKGOTLVQVFTVKHEQNTDCGGYVKLPFNSLDQTDHGDSEYNIMFGPDIC 120  
Db 78 SASPEPSNKGOTLVQVFTVKHEQNTDCGGYVKLPFNSLDQTDHGDSEYNIMFGPDIC 137  
QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 180  
Db 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVESGSL 197

RESULT 13  
ADD22407  
ID ADD22407 standard; protein; 417 AA.  
XX  
XX ADD22407;  
XX  
XX 15-JAN-2004 (first entry)  
DT  
XX HLA-B46 T cell recognised tumour antigenic polypeptide, SEQ No 57.  
DE  
XX tumour antigenic peptide; cancer; vaccine; cytostatic; cytotoxic T cell;  
KW colon; mouth; lung; prostatic; gynecological; human.  
XX  
XX Homo sapiens.  
XX  
XX JP2003111595-A.  
XX  
XX 15-APR-2003.  
XX  
XX 24-JUN-2002; 2002JP-00183603.  
XX  
XX 25-JUN-2001; 2001JP-00191974.  
XX  
XX (ITOY/) ITO Y.  
XX  
XX WPI; 2003-611129/58.  
XX  
XX Novel tumor antigenic peptide or polypeptide useful for inducing  
PT cytotoxic T cells or for treating cancer such as colon, mouth, lung,  
PT prostatic or gynecological cancer.  
XX  
XX Claim 2; SEQ ID NO 57; 98pp; Japanese.  
XX  
XX The invention relates to a novel tumour antigenic peptide or polypeptide  
CC comprising a sequence selected from 99 sequences fully defined in the  
CC specification. The tumour antigenic peptide or polypeptide comprises a

CC sequence selected from 99 sequences fully defined in the specification,  
 CC where the tumour antigenic peptide preferably has a sequence of Glu-Pro-  
 CC Pro-Leu-Ser-Gln-Glu-Thr-Phe, and the polypeptide preferably has a  
 CC sequence comprising 393 amino acids fully defined in the specification.  
 CC The invention further provides a cancer vaccine comprising a tumour  
 CC antigenic peptide or polypeptide, which has cytostatic activity. A tumour  
 CC antigenic peptide, polypeptide, its encoding polynucleotide, a  
 CC hybridising polynucleotide, a recombinant vector containing the  
 CC polynucleotide, a host transformed with the vector or an antibody are  
 CC useful for screening for compounds that interact with the tumour  
 CC antigenic peptide, the polypeptide or its encoding polynucleotide and  
 CC increases the expression of the tumour antigenic peptide, the polypeptide  
 CC or polynucleotide. The tumour antigenic peptide or the polypeptide is  
 CC useful for inducing cytotoxic T cells. The tumour antigenic peptide is  
 CC vaccine is useful for treating cancer such as colon, mouth, lung,  
 CC prostatic or gynecological cancer. The invention also provides a  
 CC pharmaceutical composition useful for treating cancer. The tumour  
 CC antigenic peptide or the polypeptide is useful as an antigen to create  
 CC antibodies. This sequence represents one of the tumour antigenic  
 CC polypeptides of the invention.

XX SQ Sequence 417 AA;

Query Match 100.0%; Score 971; DB 7; Length 417;  
 Best Local Similarity 100.0%; Pred No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKRYGDEKDKGLQTSQDARFYAL 60  
 DB 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKRYGDEKDKGLQTSQDARFYAL 77  
 QY 61 SASFEPFSNKGQTLVVQFTVKHEQNTDCGGVYVKLPNSLDQTDHGDSEYNIMFGPDIC 120  
 DB 78 SASFEPFSNKGQTLVVQFTVKHEQNTDCGGVYVKLPNSLDQTDHGDSEYNIMFGPDIC 137  
 QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLYTLVLRPDNTYEVKIDNSQVESGSLE 180  
 DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLYTLVLRPDNTYEVKIDNSQVESGSLE 197

RESULT 14

ADE56308  
 ID ADE56308 standard; protein; 417 AA.

XX ADE56308;

XX 29-JAN-2004 (first entry)

XX Human Protein NP\_004334, SEQ ID NO 2160.

XX Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

XX 14-AUG-2002; 2002WO-US025765.

XX 14-AUG-2001; 2001US-0312147P.

XX 01-NOV-2001; 2001US-0346382P.

XX 26-NOV-2001; 2001US-0333347P.

XX (GHEO ) GEN HOSPITAL CORP.

XX (FARB ) BAYER AG.

XX Woolf C, D'urso D, Befort K, Costigan M;

XX WPI; 2003-268312/26.

XX GENBANK; NP\_004334.

XX New composition comprising two or more isolated polypeptides, useful for  
 PT preparing a medicament for treating pain in an animal.

XX Claim 1; Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
 CC or human polynucleotides or a polynucleotide which represents a fragment,  
 CC derivative or allelic variation of the nucleic acid sequence. Also  
 CC claimed are a vector comprising the novel polynucleotide, a host cell  
 CC comprising the vector, a method for identifying a nucleotide sequence  
 CC which is differentially regulated in an animal subjected to pain and a  
 CC kit to perform the method, an array, a method for identifying an agent  
 CC that increases or decreases the expression of the polynucleotide sequence  
 CC that is differentially expressed in neuronal tissue of a first animal  
 CC subjected to pain, a method for identifying a compound which regulates  
 CC the expression of a polynucleotide sequence which is differentially  
 CC expressed in an animal subjected to pain, a method for identifying a  
 CC compound that regulates the activity of one or more of the  
 CC polynucleotides, a method for producing a pharmaceutical composition, a  
 CC method for identifying a compound or small molecule that regulates the  
 CC activity in an animal of one or more of the polypeptides given in the  
 CC specification, a method for identifying a compound useful in treating  
 CC pain and a pharmaceutical composition comprising the one or more  
 CC polypeptides or their antibodies. The polynucleotide or the compound that  
 CC modulates its activity is useful for preparing a medicament for treating  
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
 CC therapy). The sequence presented is a human protein (shown in Table 2 of  
 CC the specification) which is differentially expressed during pain. Note:  
 CC The sequence data for this patent did not form part of the printed  
 CC specification, but was obtained in electronic form directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 417 AA;

Query Match 100.0%; Score 971; DB 7; Length 417;  
 Best Local Similarity 100.0%; Pred No. 1.4e-98;  
 Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKRYGDEKDKGLQTSQDARFYAL 60  
 DB 18 EPAVYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKRYGDEKDKGLQTSQDARFYAL 77  
 QY 61 SASFEPFSNKGQTLVVQFTVKHEQNTDCGGVYVKLPNSLDQTDHGDSEYNIMFGPDIC 120  
 DB 78 SASFEPFSNKGQTLVVQFTVKHEQNTDCGGVYVKLPNSLDQTDHGDSEYNIMFGPDIC 137  
 QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLYTLVLRPDNTYEVKIDNSQVESGSLE 180  
 DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLYTLVLRPDNTYEVKIDNSQVESGSLE 197

RESULT 15

ADE56312

ID ADE56312 standard; protein; 417 AA.

XX ADE56312;

XX 29-JAN-2004 (first entry)

XX Human Protein NP\_004334, SEQ ID NO 2164.

XX Human; pain; neuronal tissue; gene therapy;

KW spinal segmental nerve injury; chronic constriction injury; CCI;

KW spared nerve injury; SNI; Chung.

XX Homo sapiens.

XX WO2003016475-A2.

XX 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.  
XX  
XX 14-AUG-2001; 2001US-0312147P.  
PR 01-NOV-2001; 2001US-0346382P.  
PR 26-NOV-2001; 2001US-0333347P.  
XX  
XX (GEHO ) GEN HOSPITAL CORP.  
PA (FARB ) BAYER AG.  
XX  
XX  
XX Woolf C, D'urso D, Belfort K, Costigan M;  
PI  
XX  
XX WPI; 2003-268312/26.  
DR GENBANK; NP\_004334.  
DR  
XX  
XX New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
PT  
XX  
XX Claim 1; Page; 1017pp; English:  
XX  
XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
XX  
SQ Sequence 417 AA;

Query Match 100.0%; Score 971; DB 7; Length 417;  
Best Local Similarity 100.0%; Pred. No. 1.4e-98;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 1 EPAYFYKEQFLDGGWTSRMIESKHSDFGKFLVSSGKFGYDEEKDGLQTSQDARFYAL 60  
Dy 18 EPAYFYKEQFLDGGWTSRMIESKHSDFGKFLVSSGKFGYDEEKDGLQTSQDARFYAL 77  
Qy 61 SASPEFSPNKGQTLVQFTVKHEQNDICGGGYVKLPFNSLDQTDHMGDSSEYNMFGPDIC 120  
Dy 78 SASPEFSPNKGQTLVQFTVKHEQNDICGGGYVKLPFNSLDQTDHMGDSSEYNMFGPDIC 137  
Qy 121 GPGTKKVVHVIENYKGNVLINKDRCXKDETHLYTLVPRDNTYEVKIDNSQVESGSL 180  
Dy 138 GPGTKKVVHVIENYKGNVLINKDRCXKDETHLYTLVPRDNTYEVKIDNSQVESGSL 197

Search completed: October 15, 2004, 13:15:38  
Job time : 159 secs

**This Page Blank (uspto)**

GenCore version 5.1.6.  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2004, 13:11:18 ; Search time 21 Seconds  
(without alignments)  
568.440 Million cell updates/sec

Title: US-09-807-148-4  
Perfect score: 971  
Sequence: 1 EPAYFKQFLDGDGWTSRW.....PNTYEVKIDNSQVSGSLE 180

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/1/iaa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/1/iaa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/1/iaa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/1/iaa/PCTUS.COMB.pep.\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	971	100.0	180	4	US-09-828-000-3
2	971	100.0	416	4	US-09-828-000-2
3	971	100.0	417	4	US-09-538-092-1023
4	971	100.0	417	4	US-09-906-393A-36
5	527	54.3	415	3	US-08-675-816-2
6	494	50.9	420	4	US-09-844-006A-2
7	340	35.0	61	4	US-09-828-000-4
8	327	33.7	61	4	US-09-828-000-8
9	290	29.9	593	1	US-08-296-362-2
10	288	29.7	592	4	US-09-538-092-1024
11	261	26.9	610	4	US-09-976-594-947
12	230	23.7	230	4	US-09-248-796A-14160
13	227.5	23.4	542	3	US-08-675-816-6
14	189	19.5	35	4	US-09-828-000-7
15	151	15.6	27	4	US-09-828-000-5
16	114.5	11.8	162	4	US-09-513-999C-4228
17	114.5	11.8	162	4	US-09-513-999C-4229
18	101	10.4	867	4	US-10-140-002-20
19	97	10.0	690	4	US-09-886-319A-36
20	96	9.9	18	4	US-09-828-000-6
21	89	9.2	660	3	US-09-181-706-8
22	89	9.2	660	3	US-09-458-791-8
23	89	9.2	660	3	US-09-459-066-8
24	89	9.2	660	4	US-09-459-065-8
25	81	8.3	261	4	US-09-252-991A-28346
26	81	8.3	307	4	US-09-810-347-5
27	81	8.3	360	4	US-09-810-347-4

28	81	8.3	372	4	US-09-810-347-2	Sequence 2, Appli
29	81	8.3	871	4	US-09-773-426A-3	Sequence 3, Appli
30	81	8.3	871	4	US-10-314-881-3	Sequence 3, Appli
31	81	8.3	871	4	US-09-495-823-3	Sequence 3, Appli
32	81	8.3	1611	4	US-09-668-673B-16	Sequence 16, Appli
33	81	8.3	1611	4	US-10-389-532-16	Sequence 16, Appli
34	77.5	8.0	642	2	US-08-245-511-48	Sequence 48, Appli
35	77.5	8.0	642	2	US-08-600-993A-48	Sequence 48, Appli
36	77.5	8.0	661	4	US-09-583-110-2948	Sequence 2948, Ap
37	77	7.9	15	3	US-08-946-026-53	Sequence 53, Appli
38	76	7.8	107	4	US-09-248-796A-14203	Sequence 14203, A
39	76	7.8	360	4	US-09-538-092-471	Sequence 471, App
40	76	7.8	360	3	US-09-157-257-8	Sequence 8, Appli
41	75.5	7.8	699	4	US-09-540-236-2443	Sequence 2443, Ap
42	75.5	7.8	867	4	US-09-668-673B-2	Sequence 2, Appli
43	75.5	7.8	867	4	US-10-389-532-2	Sequence 2, Appli
44	75	7.7	591	3	US-08-931-408-4	Sequence 4, Appli
45	75	7.7	591	3	US-09-432-473-4	Sequence 4, Appli

## ALIGNMENTS

RESULT 1  
US-09-828-000-3  
; Sequence 3, Application US/09828000  
; Patent No. 6596690  
; GENERAL INFORMATION:  
; APPLICANT: Government of the United States of America  
; TITLE OF INVENTION: Vasostatin as Marrow Protectant  
; FILE REFERENCE: 4239-55414  
; CURRENT APPLICATION NUMBER: US/09/828,000  
; CURRENT FILING DATE: 2001-04-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 3  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Vasostatin  
US-09-828-000-3

Query Match 100.0%; Score 971; DB 4; Length 180;  
Best Local Similarity 100.0%; Pred. No. 8.5e-102;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	EPAYFKQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL	60
DB	1	EPAYFKQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDEEKDKGLQTSQDARFYAL	60
QY	61	SASFPFSGKGTLLVQFTVKHEQNI DCGGCVKLFPNSLDQTMHGDSEYNIMEGPDIC	120
DB	61	SASFPFSGKGTLLVQFTVKHEQNI DCGGCVKLFPNSLDQTMHGDSEYNIMEGPDIC	120
QY	121	FGGTTKVVHVFNYKGNVLIINKDKDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE	180
DB	121	FGGTTKVVHVFNYKGNVLIINKDKDEFTHTLYTLVIRPDNTYEVKIDNSQVSGSLE	180

RESULT 2  
US-09-828-000-2  
; Sequence 2, Application US/09828000  
; Patent No. 6596690  
; GENERAL INFORMATION:  
; APPLICANT: Government of the United States of America  
; TITLE OF INVENTION: Vasostatin as Marrow Protectant  
; FILE REFERENCE: 4239-55414  
; CURRENT APPLICATION NUMBER: US/09/828,000  
; CURRENT FILING DATE: 2001-04-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 416  
; TYPE: PRT

; ORGANISM: Calreticulin  
US-09-828-000-2

Query Match 100.0%; Score 971; DB 4; Length 416;  
Best Local Similarity 100.0%; Pred. No. 2.8e-101;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVPEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVPEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEPFSNKGQTLVVQFTVKHEQNDICGGYVVKLFPNSLDQTMHGDSEYNIIMFGPDIC 120  
DB 78 SASPEPFSNKGQTLVVQFTVKHEQNDICGGYVVKLFPNSLDQTMHGDSEYNIIMFGPDIC 137

QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSSGSL 180  
DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSSGSL 197

RESULT 3  
US-09-538-092-1023  
; Sequence 1023, Application US/09538092  
; Patent No. 6753314  
; GENERAL INFORMATION:  
; APPLICANT: Glot, Loic  
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
; FILE REFERENCE: 15966-542  
; CURRENT APPLICATION NUMBER: US/09/538, 092  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: 60/127,352  
; PRIOR FILING DATE: 1999-04-01  
; PRIOR APPLICATION NUMBER: 60/178,965  
; PRIOR FILING DATE: 2000-02-01  
; NUMBER OF SEQ ID NOS: 1387  
; SOFTWARE: CuraPatSeqFormatter Version 0.9  
; SEQ ID NO 1023  
; LENGTH: 417  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (0)-(0)  
; OTHER INFORMATION: Polypeptide Accession Number P27797

US-09-538-092-1023

Query Match 100.0%; Score 971; DB 4; Length 417;  
Best Local Similarity 100.0%; Pred. No. 2.9e-101;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVPEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVPEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEPFSNKGQTLVVQFTVKHEQNDICGGYVVKLFPNSLDQTMHGDSEYNIIMFGPDIC 120  
DB 78 SASPEPFSNKGQTLVVQFTVKHEQNDICGGYVVKLFPNSLDQTMHGDSEYNIIMFGPDIC 137

QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSSGSL 180  
DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSSGSL 197

RESULT 4  
US-09-906-393A-36  
; Sequence 36, Application US/09906393A  
; Patent No. 6780984  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Zhou  
; APPLICANT: Xiao, Wuhan  
; TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED  
; FILE REFERENCE: 1720-1-001CIP

; CURRENT APPLICATION NUMBER: US/09/906,393A  
; CURRENT FILING DATE: 2001-07-16  
; PRIOR APPLICATION NUMBER: 60/218,761  
; PRIOR FILING DATE: 2000-07-17  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 36  
; LENGTH: 417  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-906-393A-36

Query Match 100.0%; Score 971; DB 4; Length 417;  
Best Local Similarity 100.0%; Pred. No. 2.9e-101;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVPEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVPEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYGDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEPFSNKGQTLVVQFTVKHEQNDICGGYVVKLFPNSLDQTMHGDSEYNIIMFGPDIC 120  
DB 78 SASPEPFSNKGQTLVVQFTVKHEQNDICGGYVVKLFPNSLDQTMHGDSEYNIIMFGPDIC 137

QY 121 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSSGSL 180  
DB 138 GPGTKKVHVIFNYKGNVNLINKDIRCKDDEFTHLTYTLIVRPDNTYEVKIDNSQVSSGSL 197

RESULT 5  
US-08-675-816-2  
; Sequence 2, Application US/08675816  
; Patent No. 6171864  
; GENERAL INFORMATION:  
; APPLICANT: Coughlan, Sean J.  
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND USES  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Seed and Berry  
; STREET: 701 Fifth Ave. Suite 6300  
; CITY: Seattle  
; STATE: Washington  
; COUNTRY: U.S.A.  
; ZIP: 98104-7092  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/675,816  
; FILING DATE: 05-JUL-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: No. 6171864tenburg, Carol  
; REGISTRATION NUMBER: 39,317  
; REFERENCE/DOCKET NUMBER: 750027.401  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (206-622-4900  
; TELEFAX: (206)-682-6031  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 415 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-675-816-2

Query Match 54.3%; Score 527; DB 3; Length 415;  
Best Local Similarity 55.3%; Pred. No. 4.8e-51;  
Matches 99; Conservative 31; Mismatches 43; Indels 6; Gaps 3;





REFERENCE/DOCKET NUMBER: 690066.401C1  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (206) 622-4900  
 TELEFAX: (206) 682-6031  
 TELEX: 3723836  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 593 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-296-362-2

Query Match 29.9%; Score 290; DB 1; Length 593;  
 Best Local Similarity 38.1%; Pred. No. 5.2e-24;  
 Matches 77; Conservative 26; Mismatches 65; Indels 34; Gaps 9;  
 QY 4 VYFKEQFLDGDGWTSGRWIESKHSD-----FGKVLSSGKFGYDDEK-----DKGLQTS 52  
 DB 70 VYFADSF--DRGTLSCWILSKAKKDDTDDEIAKY---DGKWEVDENKTKLPDGLGLVLM 124  
 QY 53 QDARFYALSASF-EPFSNKGOTLVQFTVKHEQNIDCGGYVKLPFN-----SLDQTDHMG 107  
 DB 125 SRKHHAISAKLNKPLFDTPKPLIVQYEVNFQNGIECGGAYVKLLSKTPPELNLDQ--FHD 182  
 QY 108 DSEYNIMFGPDICGPGTKKVVHIFNYKGNVLINKDIRCK-----DDEFTHLTYL 157  
 DB 183 KTYPTIMFGPKCGE-DYKLFHIFRHKPKTGYVEEKHAKRPDADLKTYFTDKKTHLYL 241  
 QY 158 IVRPNTYEVKIDNSQVESGSL 179  
 DB 242 ILNPDNSFEILDQSVVNSGNL 263

## RESULT 10

US-09-538-092-1024  
 Sequence 1024; Application US/09538092  
 Patent No. 6753314  
 GENERAL INFORMATION:  
 APPLICANT: Glot, Loic  
 APPLICANT: Mansfield, Traci A.  
 TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same  
 FILE REFERENCE: 15966-542  
 CURRENT APPLICATION NUMBER: US/09/538,092  
 CURRENT FILING DATE: 2000-03-29  
 PRIOR APPLICATION NUMBER: 60/127,352  
 PRIOR FILING DATE: 1999-04-01  
 PRIOR APPLICATION NUMBER: 60/178,965  
 PRIOR FILING DATE: 2000-02-01  
 NUMBER OF SEQ ID NOS: 1387  
 SOFTWARE: CurapacSeqformatter Version 0.9  
 SEQ ID NO 1024  
 LENGTH: 592  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 LOCATION: (0)...(0)  
 OTHER INFORMATION: Polypeptide Accession Number P27824  
 US-09-538-092-1024

Query Match 29.7%; Score 288; DB 4; Length 592;  
 Best Local Similarity 38.1%; Pred. No. 8.7e-24;  
 Matches 77; Conservative 26; Mismatches 65; Indels 34; Gaps 9;  
 QY 4 VYFKEQFLDGDGWTSGRWIESKHSD-----FGKVLSSGKFGYDDEK-----DKGLQTS 52  
 DB 69 VYFADSF--DRGTLSCWILSKAKKDDTDDEIAKY---DGKWEVDENKTKLPDGLGLVLM 123  
 QY 53 QDARFYALSASF-EPFSNKGOTLVQFTVKHEQNIDCGGYVKLPFN-----SLDQTDHMG 107  
 DB 124 SRKHHAISAKLNKPLFDTPKPLIVQYEVNFQNGIECGGAYVKLLSKTPPELNLDQ--FHD 181

QY 108 DSEYNIMFGPDICGPGTKKVVHIFNYKGNVLINKDIRCK-----DDEFTHLTYL 157  
 DB 182 KTYPTIMFGPKCGE-DYKLFHIFRHKPKTGYVEEKHAKRPDADLKTYFTDKKTHLYL 240  
 QY 158 IVRPNTYEVKIDNSQVESGSL 179  
 DB 241 ILNPDNSFEILDQSVVNSGNL 262

## RESULT 11

US-09-976-594-947  
 Sequence 947; Application US/09976594  
 Patent No. 6673549  
 GENERAL INFORMATION:  
 APPLICANT: Furness, Michael  
 APPLICANT: Buchbinder, Jenny  
 TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS  
 FILE REFERENCE: PA-0041 US  
 CURRENT APPLICATION NUMBER: US/09/976,594  
 CURRENT FILING DATE: 2001-10-12  
 PRIOR APPLICATION NUMBER: 60/240,409  
 PRIOR FILING DATE: 2000-10-12  
 NUMBER OF SEQ ID NOS: 1143  
 SOFTWARE: PERL Program  
 SEQ ID NO 947  
 LENGTH: 610  
 TYPE: PRT  
 ORGANISM: Homo sapiens  
 FEATURE:  
 NAME/KEY: misc feature  
 OTHER INFORMATION: Incyte ID No. 6673549 3876162CD1  
 US-09-976-594-947

Query Match 26.9%; Score 261; DB 4; Length 610;  
 Best Local Similarity 34.5%; Pred. No. 1e-20;  
 Matches 68; Conservative 33; Mismatches 72; Indels 24; Gaps 7;

QY 4 VYFKEQFLDGDGWTSGRWIESKHSD--FGKVLSSGKFGYDDEK-----DKGLQTSQDA 55  
 DB 60 VYFAETF--DSGRLAGWLSKAKKDDMDDEISIDGRWEIEELKENQVPGDGLVKSR 117  
 QY 56 RFVALSASF-EPFSNKGOTLVQFTVKHEQNIDCGGYVKLPFNLD--QTMHGDSEVN 112  
 DB 118 KHAISAVLAKPFIADKPLIVQYEVNFQDGGAYIKLADTDLLENFYDKTSYI 177  
 QY 113 IMFPGDICGPGTKKVVHIFNYKGNVLINKDIRCK-----DDEFTHLTYLIVRPD 162  
 DB 178 IMFPGPKCGE-DYKLFHIFRHKPKTGYVEEKHAKPPVDLKKKFTDRKTHLYLVMNP 236  
 QY 163 NTYEVKIDNSQVESGSL 179  
 DB 237 DTFEVLVDQTVVNSGNL 253

## RESULT 12

US-09-248-796A-14160  
 Sequence 14160; Application US/09248796A  
 Patent No. 6747137  
 GENERAL INFORMATION:  
 APPLICANT: Keith Weinstein et al  
 TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN  
 FILE REFERENCE: 107196.132  
 CURRENT APPLICATION NUMBER: US/09/248,796A  
 CURRENT FILING DATE: 1999-02-12  
 PRIOR APPLICATION NUMBER: US 60/074,725  
 PRIOR FILING DATE: 1998-02-13  
 PRIOR APPLICATION NUMBER: US 60/096,409  
 PRIOR FILING DATE: 1998-08-13  
 NUMBER OF SEQ ID NOS: 28208  
 SEQ ID NO 14160  
 LENGTH: 230  
 TYPE: PRT

```

; ORGANISM: Candida albicans
US-09-248-796A-14160

Query Match      23.7%; Score 230; DB 4; Length 230;
Best Local Similarity 32.0%; Pred.No. 8.1e-18;
Matches 66; Conservative 39; Mismatches 69; Indels 32; Gaps 10;

QY 1 EPAYVPEQF-----LDGDGWTSRWIESKHKSDFGK--FVLSSGKF-----YGDDEKD 46
DB 27 DPASVP-EQDYPSLNSPWQ---VSTAKFDEGRDEIVRSGEWKIESSTSKYPGLEGD 82
QY 47 KGLQTSQDARFYALS-----ASFEPFSNKGQTLVVQFTVKHEQNIDCGGYVKLPFNS 99
DB 83 LGLVMKSRAGHYAISYKLPHEVTNTNPNNNKTQDLVLQYEVKLGGLTCGGAYIKLLDSS 142
QY 100 -LQCTDMHGDSSEYNIMFGPDI CGGTGKVHVIFNYKGNVLI-NKDIR----CKDDPETH 153
DB 143 PSGYKFNSSETPQIMFGPDVCG-SENKIHFIIRKLPNGAIEBKHLKHKPMARTNELTN 201
QY 154 LYLTLIVRPDNTYEVKIDNSQVESGSL 179
DB 202 LYLTLIKFNSDFEIRINGTVAKAGNL 227

RESULT 13
US-08-675-816-6
; Sequence 6, Application US/08675816
; Patent No. 6171864
; GENERAL INFORMATION:
; APPLICANT: Coughlan, Sean J.
; APPLICANT: Winfrey, Jr., Ron J.
; TITLE OF INVENTION: CALRETICULIN AND CALNEXIN GENES AND PROMOTER REGIONS AND USES
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seed and Berry
; STREET: 701 Fifth Ave. Suite 6300
; CITY: Seattle
; STATE: Washington
; COUNTRY: U.S.A.
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/675,816
; FILING DATE: 05-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6171864tenburg, Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 750027.401
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206)-622-4900
; TELEFAX: (206)-682-6031
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 542 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-675-816-6

Query Match      23.4%; Score 227.5; DB 3; Length 542;
Best Local Similarity 36.0%; Pred. No. 5.4e-17;
Matches 64; Conservative 24; Mismatches 73; Indels 17; Gaps 10;

QY 6 FKEQFLDGDGWTSRWIESKHKSDFGKFVLSGKPFYGDDEKDKGLQTSQDARFYALSASF- 64
DB 35 FYESF--EDSFEGRWTLSA-KDDY-KGWKHKAKSEGHDE--YGLLVSEKARKYAIKVELD 88
QY 65 EPPFNKGQTLVVQFTVRHEQNIDCGGYVK-LFPNSLDQT--DMHGDSSEYNIMFGPDTCG 121

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 15, 2004, 13:12:37 ; Search time 49 Seconds  
(without alignments)  
1187.702 Million cell updates/sec

Title: US-09-807-148-4  
Perfect score: 971  
Sequence: 1 EPAYFKQFLDGDGWTSRW.....PDNTYEVKIDNSQVSGSLE 180

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1360919 seqs, 32331874 residues

Total number of hits satisfying chosen parameters: 1360919

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/prodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/prodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/prodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/prodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/prodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/prodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/prodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/prodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
10: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
11: /cgn2\_6/prodata/1/pubpaa/US09\_PUBCOMB.pep.\*  
12: /cgn2\_6/prodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
14: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
15: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
16: /cgn2\_6/prodata/1/pubpaa/US10\_PUBCOMB.pep.\*  
17: /cgn2\_6/prodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
18: /cgn2\_6/prodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
19: /cgn2\_6/prodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	971	100.0	180	10	US-09-828-000-3
2	971	100.0	180	14	US-10-405-588-3
3	971	100.0	416	10	US-09-828-000-2
4	971	100.0	416	14	US-10-405-588-2
5	971	100.0	417	10	US-09-906-393A-36
6	971	100.0	417	14	US-10-161-959-29
7	971	100.0	417	14	US-10-367-093-14
8	933	96.1	416	14	US-10-316-253-2
9	933	96.1	416	14	US-10-316-253-4
10	933	96.1	416	14	US-10-316-253-6
11	631	65.0	390	14	US-10-161-927-46
12	619.5	63.8	395	14	US-10-369-493-6343
13	616.5	63.5	177	15	US-10-264-049-3546
14	584	60.1	210	16	US-10-767-701-55877

15	536	55.2	435	15	US-10-425-114-70304	Sequence 70304, A
16	530	54.6	421	16	US-10-767-701-46544	Sequence 46544, A
17	529	54.5	442	16	US-10-437-963-114914	Sequence 114914, A
18	529	54.5	667	16	US-10-437-963-114860	Sequence 114860, A
19	527	54.3	424	16	US-10-437-963-201427	Sequence 201427, A
20	526	54.2	424	16	US-10-437-963-201420	Sequence 201420, A
21	523.5	53.9	388	15	US-10-425-114-46405	Sequence 46405, A
22	523.5	53.9	391	15	US-10-425-114-69619	Sequence 69619, A
23	523.5	53.9	442	15	US-10-425-114-46869	Sequence 46869, A
24	517	53.2	420	15	US-10-424-599-190958	Sequence 190958, A
25	512	52.7	431	16	US-10-767-701-44015	Sequence 44015, A
26	503	51.8	391	15	US-10-424-599-190956	Sequence 190956, A
27	503	51.8	391	15	US-10-425-114-46228	Sequence 46228, A
28	503	51.8	391	15	US-10-425-114-55395	Sequence 55395, A
29	494	50.9	420	9	US-09-844-006A-2	Sequence 2, Appli
30	489	50.4	425	16	US-10-437-963-122222	Sequence 122222, A
31	483	49.7	290	15	US-10-425-114-44768	Sequence 44768, A
32	483	49.7	434	15	US-10-425-114-70140	Sequence 70140, A
33	483	49.7	439	15	US-10-425-114-65495	Sequence 65495, A
34	479	49.3	422	15	US-10-424-599-277445	Sequence 277445, A
35	476	49.0	279	16	US-10-767-701-39041	Sequence 39041, A
36	468	48.2	422	16	US-10-767-701-45264	Sequence 45264, A
37	467	48.1	247	15	US-10-424-599-153809	Sequence 153809, A
38	466	48.0	372	15	US-10-425-114-44621	Sequence 44621, A
39	466	48.0	372	15	US-10-425-114-54997	Sequence 54997, A
40	466	48.0	418	15	US-10-424-599-153808	Sequence 153808, A
41	465.5	47.9	993	16	US-10-437-963-114903	Sequence 114903, A
42	460	47.4	434	15	US-10-425-114-57598	Sequence 57598, A
43	438	45.1	161	16	US-10-767-701-61035	Sequence 61035, A
44	421	43.4	346	15	US-10-425-114-48930	Sequence 48930, A
45	340	35.0	61	10	US-09-828-000-4	Sequence 4, Appli

## ALIGNMENTS

### RESULT 1

US-09-828-000-3  
; Sequence 3, Application US/09828000  
; Publication No. US20030078198A1  
; GENERAL INFORMATION:  
; APPLICANT: Government of the United States of America  
; TITLE OF INVENTION: Vasostatin as Marrow Protectant  
; FILE REFERENCE: 4239-55414  
; CURRENT APPLICATION NUMBER: US/09/828,000  
; CURRENT FILING DATE: 2001-04-06  
; NUMBER OF SEQ ID NOS: 8  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 3  
; LENGTH: 180  
; TYPE: PRT  
; ORGANISM: Vasostatin  
US-09-828-000-3

Query Match 100.0%; Score 971; DB 10; Length 180;

Best Local Similarity 100.0%; Pred. No. 1.8e-95; Mismatches 0; Indels 0; Gaps 0;

Qy	1	EPAYFKQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDGEEDKGLQTSQDARFYAL	60
Db	1	EPAYFKQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDGEEDKGLQTSQDARFYAL	60
Qy	61	SASFPFNNKGQTLVVOFTVKHEQINDCGGGYKLFNPSLDQTDHMGDSEYNIMEGPDIC	120
Db	61	SASFPFNNKGQTLVVOFTVKHEQINDCGGGYKLFNPSLDQTDHMGDSEYNIMEGPDIC	120
Qy	121	GPCTKKVHVFNYKGNVLINKDKRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE	180
Db	121	GPCTKKVHVFNYKGNVLINKDKRCKDDETHLYTLIVRPDNTYEVKIDNSQVSGSLE	180

### RESULT 2

US-10-405-588-3

	Query Match	100.0%	Score 971;	DB 14;	Length 416;
	Best Local Similarity	100.0%;	Pred. No. 5.6e-95;		
	Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	BPAYFKSQFLDGGWTSRMIESKHKSDFGKFVLSGKFGYGDDEKDKGLQTSQDARFVAL	60		
Db	18	BPAYFKSQFLDGGWTSRMIESKHKSDFGKFVLSGKFGYGDDEKDKGLQTSQDARFVAL	77		
Qy	61	SASFEPPFNKQQTLLVQFTVKHQEINDCGGGYVKLPFNSLDQTDHXGDSYNNIMFGPDIC	120		
Db	78	SASFEPPFNKQQTLLVQFTVKHQEINDCGGGYVKLPFNSLDQTDHXGDSYNNIMFGPDIC	137		
Qy	121	SPGTKKVHVIFNYGKGNVLINKIRCKDBEFTHLYTLIVRPNTVEVKIDNSQVTSGLSLE	180		
Db	138	SPGTKKVHVIFNYGKGNVLINKIRCKDBEFTHLYTLIVRPNTVEVKIDNSQVTSGLSLE	197		

```

RESULT 5
US-09-906-393A-36
/ Sequence 36, Application US/09906393A
/ Publication No. US20030039970A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Zhou
/ APPLICANT: Xiao, Wuhan
/ TITLE OF INVENTION: METHOD OF PROGNOSING CANCER AND THE PROTEINS INVOLVED
/ FILE REFERENCE: 1720-1-001CIP
/ CURRENT APPLICATION NUMBER: US/09/906,393A
/ CURRENT FILING DATE: 2001-07-16
/ PRIOR APPLICATION NUMBER: 60/218,761
/ PRIOR FILING DATE: 2000-07-17
/ NUMBER OF SEQ ID NOS: 36
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 36
/ LENGTH: 417
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-09-906-393A-36

```

	Query Match	100.0%;	Score 971;	DB 10;	Length 417;
	Best Local Similarity	100.0%;	Pred. No. 5.6e-95;		
	Matches 180;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	EPAYFYKEQPLDGDGWTSRWIESKHKXSDRGKFKVLSSGKPYGDEEXDKGLQTSQDARFYAL	60		
Db	18	EPAYFYKEQPLDGDGWTSRWIESKHKXSDRGKFKVLSSGKPYGDEEXDKGLQTSQDARFYAL	77		
Qy	61	SASPEPSNKGQTLVQVFTVKHEQNTDCGGGVYKLFPNSLDOTDMHGDSSEYNIMFGPDIC	120		
Db	78	SASPEPSNKGQTLVQVFTVKHEQNTDCGGGVYKLFPNSLDOTDMHGDSSEYNIMFGPDIC	137		
Qy	121	GPGTKVHVIFNYKGNVNLINKDIRCKDDEFTHLVTLVRPNTVEYKIDNSQVESGSLE	180		
Db	138	GPGTKVHVIFNYKGNVNLINKDIRCKDDEFTHLVTLVRPNTVEYKIDNSQVESGSLE	197		

## RESULT 6

US-10-161-959-29  
; Sequence 29, Application US/10161959  
; Publication No. US20030095748A1  
; GENERAL INFORMATION:  
; APPLICANT: Holoshitz, Joseph  
; TITLE OF INVENTION: Methods and Compositions for the Treatment of Diseases Associated  
; WITH SIGNAL TRANSDUCTION ABERRATIONS  
; FILE REFERENCE: US-07135  
; CURRENT APPLICATION NUMBER: US/10/161,959  
; CURRENT FILING DATE: 2002-10-01  
; PRIOR APPLICATION NUMBER: 60/295,691  
; PRIOR FILING DATE: 2001-06-04  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 29  
; LENGTH: 417  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Synthetic  
US-10-161-959-29

Query Match 100.0%; Score 971; DB 14; Length 417;  
Best Local Similarity 100.0%; Pred. No. 5.6e-95;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
DB 78 SASPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 137

QY 121 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180  
DB 138 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 7  
US-10-367-093-14  
; Sequence 14, Application US/10367093  
; Publication No. US20030216315A1  
; GENERAL INFORMATION:  
; APPLICANT: Duke University  
; APPLICANT: Nicchitta, Chris  
; APPLICANT: Baker-LePain, Julie  
; TITLE OF INVENTION: MODULATION OF IMMUNE RESPONSE BY NON-PEPTIDE BINDING STRESS RESP  
; TITLE OF INVENTION: POLYPEPTIDES  
; FILE REFERENCE: 180/145  
; CURRENT APPLICATION NUMBER: US/10/367,093  
; CURRENT FILING DATE: 2003-02-13  
; NUMBER OF SEQ ID NOS: 27  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 14  
; LENGTH: 417  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-367-093-14

Query Match 100.0%; Score 971; DB 14; Length 417;  
Best Local Similarity 100.0%; Pred. No. 5.6e-95;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
DB 78 SASPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 137

US-10-161-959-29

Query Match 100.0%; Score 971; DB 14; Length 417;  
Best Local Similarity 100.0%; Pred. No. 5.6e-95;  
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
DB 18 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
DB 78 SASPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 137

QY 121 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180  
DB 138 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 8  
US-10-316-253-2  
; Sequence 2, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 2  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-316-253-2

Query Match 96.1%; Score 933; DB 14; Length 416;  
Best Local Similarity 94.4%; Pred. No. 6.4e-91;  
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

QY 1 EPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 60  
DB 18 DPAYVFEQFLDGGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFYAL 77

QY 61 SASPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 120  
DB 78 SARPEFSPNKGQTLVQVFTVKHEQNDICCGGYVKLPNSLDQTMHGDSEYNIMFGPDIC 137

QY 121 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 180  
DB 138 GPCTKKVHVIENYKGNVLINKDIRCKDDEFTHTLYTLIVRPDNTYEVKIDNSQVSGSLE 197

RESULT 9  
US-10-316-253-4  
; Sequence 4, Application US/10316253  
; Publication No. US20030162706A1  
; GENERAL INFORMATION:  
; APPLICANT: The Procter & Gamble Company  
; APPLICANT: Peters, Kevin  
; APPLICANT: Thompson, Larry  
; APPLICANT: Wang, Feng  
; APPLICANT: Greis, Kenneth  
; TITLE OF INVENTION: Angiogenesis Modulating Proteins  
; FILE REFERENCE: 8865M  
; CURRENT APPLICATION NUMBER: US/10/316,253  
; CURRENT FILING DATE: 2002-12-10  
; PRIOR APPLICATION NUMBER: US 60/355,295  
; PRIOR FILING DATE: 2002-02-08  
; NUMBER OF SEQ ID NOS: 308  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 4  
; LENGTH: 416  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-316-253-4

Query Match 96.1%; Score 933; DB 14; Length 416;  
Best Local Similarity 94.4%; Pred. No. 6.4e-91;  
Matches 170; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

[illegible]

```

RESULT 10
US-10-316-253-6
; Sequence 6, Application US/10316253
; Publication NO. US20030162706A1
; GENERAL INFORMATION:
; APPLICANT: The Procter & Gamble Company
; APPLICANT: Peters, Kevin
; APPLICANT: Thompson, Larry
; APPLICANT: Wang, Feng
; APPLICANT: Greis, Kenneth
; TITLE OF INVENTION: Angiogenesis Modulating Proteins
; FILE REFERENCE: 8865M
; CURRENT APPLICATION NUMBER: US/10/316,253
; CURRENT FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: US 60/355,295
; PRIOR FILING DATE: 2002-02-08
; NUMBER OF SEQ ID NOS: 308
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 416
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-316-253-6

```

Query Match	96.1%	Score 933	DB 14	Length 416
Best Local Similarity	94.4%	Pred. No. 6.4e-91		
Matches 170	Conservative 5	Mismatches 5	Indels 0	Gaps 0
Qy	1	EPAVYFKQFIDGCGWTSRWLESKHSKGFDFKGVLSGKFGYGDDEKDKG:QTSQDARFVAL	60	
Db	18	DPATYFKQFIDGCGAWTNRWVESKHSKGFDFKGVLSGKFGYGDDEKDKG:QTSQDARFVAL	77	
Qy	61	SASFEPFNKGQTLVQVQTVTAEHQNICGGGYVKLFPSLDDQTDHGDSEYNIIMFGPIC	120	
Db	78	SARFEPFNKGQTLVQVQTVTAEHQNICGGGYVKLFPGLDQDKMGDSEYNIIMFGPIC	137	
Qy	121	GGTKKKVAVINYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSSGSL	180	
Db	138	GGTKKKVAVINYKGNVLINKDIRCKDDFTHLYTLIVRPDNTYEVKIDNSQVSSGSL	197	

RESULT 11  
US-10-161-927-46  
; Sequence 46, Application US/10161927  
; Publication No. US20030235821A1  
; GENERAL INFORMATION:  
; APPLICANT: Zerhusen, Bryan D.  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Shenoy, Suresh G.  
; APPLICANT: Miller, Charles E.  
; APPLICANT: Hjal, Tord  
; APPLICANT: Gerlach, Valerie L.  
; APPLICANT: Baumgartner, Jason C.  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Gangolli, Esha A.  
; APPLICANT: Vernet, Corine  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Li, Li



```

; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 6343
; LENGTH: 395
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-10-369-493-6343

Query Match      63.8%; Score 619.5; DB 14; Length 395;
Best Local Similarity 63.5%; Pred. No. 1.9e-57;
Matches 113; Conservative 26; Mismatches 36; Indels 3; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFVALSAS 63
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 17 VYFKEFNDA-SWEKRWQSGKHDDFGAFKLSAGKFFDVESRDQGIQTSQDAKFYSRAAK 75
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 64 FE-PFSGKQTLVQVFTVKHEQNIDCGGGYVKLFPSNLDQTMHGDSEYNIMFGPDICGP 122
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 76 FDKDFSNKGTLLVIQVTKHEQIDCGGGYVKMRADADLGDPHGTPYVNMFGPDICGP 135
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 123 GTKKVHVIFNYKGNVLINKDIRCKDDETHLYTLVRDNTYEVKIDNSQVESGSL 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 136 -TRRVHVLNKGKGLIKKEITCKSDDELTHLYTLNDSNTYEVKIDGESAQTSLSLE 192
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 13
US-10-264-049-3546
; Sequence 3546, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133p1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 3546
; LENGTH: 177
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (138)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-049-3546

Query Match      63.5%; Score 616.5; DB 15; Length 177;
Best Local Similarity 84.6%; Pred. No. 1.4e-57;
Matches 121; Conservative 3; Mismatches 18; Indels 1; Gaps 1;

QY 1 EPAYVFKQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFVAL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 29 EPAYVFKQFLDGDGWTSRWIESKHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFVAL 88
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 SASPEPFSNKGQTLVQVFTVKHEQNIDCGGGYVKLFPSNLDQTMHGDSEYNIMFGPDIC 119
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 89 SASPEPFSNKGQTLVQVFTVKHEQNIDCGGGYVKLFVIVWTRQTCGNSKYNIMVGRHL 148
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 120 CGPCTKKVHVIFNYKGNVLINK 142
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 149 XALAPKKVHVIFNYKGNGLINK 171
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 14
US-10-767-701-55877

```

```

; Sequence 55877, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5335)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 55877
; LENGTH: 210
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: 30933722.pep
US-10-767-701-55877

Query Match      60.1%; Score 584; DB 16; Length 210;
Best Local Similarity 60.0%; Pred. No. 5.1e-54;
Matches 108; Conservative 30; Mismatches 36; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDG-WTSRWIES--XHKSDFGKFLVSSGKFGYDDEKDKGLQTSQDARFVAL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 VYFKETF---DGKWDRTWTESEWKEAGTRGKFAVYTAGKFYNDDEADKGLQTTEDYRYITI 82
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 SASPEPFSNKGQTLVQVFTVKHEQNIDCGGGYVKLFPSNLDQTMHGDSEYNIMFGPDIC 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 SADHEEFSNKGKTLVLQYSLNQQKLDGCGGYLKFPPAGVDKKTGLNGDSKYNIMFGPDIC 142
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 121 GPGTKKHVIFNYKGNVLINKDIRCKDDETHLYTLVRDNTYEVKIDNSQVESGSL 180
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 143 GTTRKHVIFHEHGNHVLKKEIPCETDETHLYTLVRDNTFEVLVDGVSATGSL 202
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

RESULT 15
US-10-425-114-70304
; Sequence 70304, Application US/10425114
; Publication No. US20040304888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70304
; LENGTH: 435
; TYPE: PRT
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMR0TE0SINTE102C08_FLI.pep
US-10-425-114-70304

Query Match      55.2%; Score 536; DB 15; Length 435;
Best Local Similarity 56.4%; Pred. No. 1.8e-48;
Matches 101; Conservative 31; Mismatches 41; Indels 6; Gaps 3;

QY 4 VYFKEQFLDGDGWTSRWIESKHKSD---FGKFLVSSGKFGYDDEKDKGLQTSQDARFVAL 60
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 41 VFFQKFP--EDWESRWVKSEWKDENWAGWNHTSGKNGDAE-DKGIQTSSEYRFVAI 97
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

QY 61 SASPEPFSNKGQTLVQVFTVKHEQNIDCGGGYVKLFPSNLDQTMHGDSEYNIMFGPDIC 120
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 98 SAEYEPFSNKGKTLVLQFSVKHEQKLDGCGGYVKLLGGVDQKFKPGGDTSYSINFGPDIC 157
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

Qy 121 GPCTKKVHVIFNYKGNVLINKDIRCKDDFTHLVTLIVRPDNTYEVKIDNSQVESGSL 179  
Db 158 GYSTKKVHTILTCKGKNHLIKKDVPCETDQLTHVYTLIIRPDATYSILIDNEEKOTGSI 216

Search completed: October 15, 2004, 13:17:09  
Job time : 49 secs